

Modelling Trajectories of Social and Non-Social Development in Infants at High Risk for
Autism Spectrum Disorder

Rachael Bedford

Doctor of Philosophy

Institute of Education

2012

Declaration

I, Rachael Bedford, confirm that all the work presented in this thesis is my own.

My PhD project is based on data from the Phase 1 cohort of the British Autism Study of Infant Siblings (BASIS). Although my focus is analysis of data from this project, I have also been involved in every stage of the BASIS project, from scheduling families to testing participants to entering and analysing data. More specifically, I tested 24-month-olds and 36-month-olds on experimental tasks using techniques such as eye-tracking and electroencephalography (EEG) as well as administering developmental and clinical assessments such as the Mullen Scales of Early Learning (*MSEL*, Mullen, 1995) and the Autism Diagnostic Observation Schedule (ADOS-G, Lord et al., 2000).

Based on the work presented in Chapters 3 and 4 of this thesis, I have published two first author articles in peer reviewed journals. The gaze following study presented in Chapter 3 is published in the *Journal of Autism and Developmental Disorders*:

Bedford, R., Elsabbagh, M., Senju, A., Gliga, T., Pickles, A., Charman, T., Johnson, M. H. and the BASIS team. (2012). Precursors to Social and Communication Difficulties in Infants At-Risk for Autism: Gaze Following and Attentional Engagement. *Journal for Autism and Developmental Disorders* DOI: 10.1007/s10803-012-1450-y

The mutual exclusivity study, in Chapter 4, is in press in the *Journal of Child Language*:

Bedford, R., Gliga, T., Frame, K., Hudry, K., Chandler, S., Johnson, M. H., Charman, T. and the BASIS team. Failure to learn from feedback underlies word learning difficulties in toddlers at-risk for autism? (in press *Journal of Child Language*)

18th August 2012

RM Bedford.

Acknowledgements

The work presented in this thesis has been supported by a Bloomsbury Colleges Scholarship to R. Bedford, a Leverhulme Early Career Fellowship to M. Elsabbagh, COST Action BM1004 to T. Charman and M. Elsabbagh, the BASIS funding consortium led by Autistica (www.basisnetwork.org), Autism Speaks (PI M. H. Johnson, Grant number 1292) and a UK Medical Research Council Programme Grant (G0701484) to M. H. Johnson. The Centre for Research in Autism and Education is supported by The Clothworkers' Foundation and Pears Foundation.

I would like to thank my two supervisors, Professor Tony Charman and Professor Mark Johnson for all their support and encouragement over the past 3 years. Without their inspiration and help I would not have been able to complete this PhD. I would also like to thank Professor Andrew Pickles and Dr. Teodora Gliga for their constant help and mentorship, and for reading drafts of my thesis. I am grateful to all the members of the Centre for Brain and Cognitive Development, Birkbeck College, and the Centre for Research in Autism and Education, in particular, Leslie Tucker, Mayada Elsabbagh, Ted Barker, Cathy Manning and Liz Pellicano, who have helped me throughout my PhD.

I am especially grateful to my parents, William and Fiona Bedford who have been there to support me every step of the way, to my brother, Tom Bedford, for proof-reading my thesis, and to Tim Smith who patiently and repeatedly encouraged me throughout the write up of my thesis.

Finally, I would like to acknowledge the many amazing families who took part in this research; without whose help and enthusiasm this project would not have been possible.

Abstract

In this thesis, multivariate statistical modelling approaches are applied to longitudinal data to examine how social and non-social abilities in infancy relate to later developmental and clinical outcomes in infants at high- and low-risk for autism spectrum disorder (ASD). ASD is a heritable developmental disorder, rarely diagnosed before 2 years, and characterised by impairments in social interaction, communication and restricted and repetitive behaviours (DSM-IV-TR, American Psychiatric Association, 2000). The application of multivariate techniques to experimental, clinical and questionnaire data from 54 high-risk infants (who have an older sibling with ASD) and 50 low-risk controls, seen at 7, 13, 24 and 36 months, enables the simultaneous modelling of multiple factors in relation to ASD outcome.

In Chapter 3, eye-tracking of gaze following behaviour demonstrated that infants who later develop ASD symptoms can correctly follow another person's eye-gaze, but by 13 months they show reduced looking to the gazed-at object compared to typically developing infants. Looking time is an index of referential understanding and was also related to children's subsequent receptive vocabulary. Chapter 4 reported that both high- and low-risk 24-month-olds can apply mutual exclusivity (the principle that each object has one name) to make a word-object mapping. However, high-risk toddlers did not use social feedback to learn the word, and this difficulty was related to their lower receptive vocabularies. Further, 13 month looking time (from the gaze following task) predicted later learning from reinforcing feedback, suggesting a degree of continuity in children's social understanding across development. Finally, Chapter 5 found that social (gaze following) and non-social (disengagement) attention independently predict ASD, and while disengagement predicts looking time early in development, the measures become de-correlated over time. The findings suggest that in order to understand the variable developmental trajectories leading to ASD, multiple risk markers over time should be analysed.

Contents

List of Tables.....	11
List of Figures.....	12
List of Acronyms.....	15
Chapter 1. Introduction: Infant Predictors of Later Outcomes in Typical and Atypical Development.....	18
<i>1.1 Introduction</i>	18
<i>1.2 Approaches for predicting later outcomes</i>	19
<i>1.3 Infant predictors of later outcomes in typical development</i>	22
<i>1.4 Predictors of later outcomes in atypical development</i>	26
<i>1.5 Predictors of the development of ASD</i>	27
<i>1.6 Social development</i>	29
<i>1.6.1 Prospective social studies</i>	34
<i>1.7 Non-social development</i>	37
<i>1.7.1 Prospective non-social studies</i>	41
<i>1.8 What comes first: social or non-social?</i>	43
<i>1.9 Conclusion</i>	45
<i>Summary of Chapter1</i>	48
Chapter 2. General Methods: An Introduction to Developmental Statistical Modelling..	49
2.1 Introduction.....	49
2.2 Participants.....	50
2.3 Overview of the data set.....	51
2.3.1 Clinical and developmental assessments.....	51
2.3.2 Experimental measures.....	54

2.4 Longitudinal data analysis.....	58
2.5 Structural equation modelling.....	62
2.6 Data analysis: the Mullen Scales of Early Learning.....	67
2.6.1 Autoregressive models.....	70
2.6.2 Growth curve models.....	73
2.6.3 Growth mixture models.....	79
2.7 Summary of the methodological approach.....	80
Summary of Chapter 2.....	81
Chapter 3. The Gaze Following Task: Eye-tracking of Gaze Following Behaviour in Infants at High Risk for Autism Spectrum Disorder.....	82
3.1 Introduction.....	82
Part 1: Precursors to Social and Communication Difficulties in Infants at High Risk for Autism: Gaze Following and Attentional Engagement.....	83
3.2 The emergence of gaze following and joint attention in typical development.....	83
3.2.1 Prospective studies of infants at high risk for ASD.....	85
3.2.2 The gaze following task.....	87
3.3 Method.....	88
3.3.1 Participants.....	88
3.3.2 Behavioural assessment and outcome groups.....	89
3.3.3 Apparatus.....	90
3.3.4 Stimuli and procedure.....	91
3.3.5 Data analysis.....	92
3.3.6 Generalised estimating equation (GEE) analysis.....	94
3.4 Results.....	95

3.4.1	<i>Seven month visit</i>	95
3.4.2	<i>Thirteen month visit</i>	97
3.5	<i>Discussion</i>	99
<i>Part 2: Prediction of 24 Month Parent-Reported Vocabulary by Early Gaze Following Behaviour</i>		105
3.6	<i>Gaze following and vocabulary</i>	105
3.6.1	<i>Links between JA and vocabulary in typical development</i>	105
3.6.2	<i>Links between JA and vocabulary in ASD</i>	106
3.6.3	<i>Infants at high risk for ASD</i>	107
3.6.4	<i>Links between gaze following and language</i>	107
3.6.5	<i>The current study</i>	108
3.7	<i>Method</i>	108
3.7.1	<i>Data analysis</i>	108
3.8	<i>Results</i>	108
3.9	<i>Discussion</i>	112
3.10	<i>Conclusion</i>	115
<i>Summary of Chapter 3</i>		117
 Chapter 4. Problems in Understanding Social Cues Over Development Underlie Word Learning Difficulties in Children at High Risk for Autism Spectrum Disorder		
		118
4.1	<i>Introduction</i>	118
<i>Part 1: The use of Mutual Exclusivity and Social Feedback to Learn Words in 24-month-olds at High Risk for Autism Spectrum Disorder</i>		119
4.2	<i>The mutual exclusivity task</i>	119
4.2.1	<i>Word learning in typical development</i>	119

4.2.2	<i>Word learning in ASD</i>	121
4.2.3	<i>Word learning in infants at high risk for ASD</i>	122
4.3	<i>Method</i>	123
4.3.1	<i>Participants</i>	123
4.3.2	<i>Stimuli</i>	123
4.3.3	<i>Procedure and design</i>	124
4.3.4	<i>Data analysis</i>	126
4.3.5	<i>Measures of language and general development</i>	127
4.4	<i>Results</i>	128
4.4.1	<i>Fast-mapping</i>	128
4.4.2	<i>Retention trials</i>	130
4.4.3	<i>Type of feedback</i>	131
4.4.4	<i>Relationship with receptive vocabulary size</i>	132
4.5	<i>Discussion</i>	133
<i>Part 2: Developmental Continuity in Problems Understanding Social Cues in Children at High Risk for Autism Spectrum Disorder</i>		138
4.6	<i>Gaze following and mutual exclusivity</i>	138
4.7	<i>Method</i>	140
4.8	<i>Results</i>	141
4.9	<i>Discussion</i>	143
<i>Summary of Chapter 4</i>		146
Chapter 5. Developmental Interactions between Social and Non-Social Attention in Infants at High Risk for Autism Spectrum Disorder		147
5.1	<i>Introduction</i>	147

<i>Part 1: Gaze Following and Visual Disengagement Independently Predict the Development of Autism Spectrum Disorder</i>	148
5.2 <i>The autistic triad</i>	148
5.2.1 <i>Cumulative and cascading models of risk</i>	150
5.3 <i>Why might visual attention be important in the development of ASD?</i>	152
5.3.1 <i>Attentional disengagement difficulties in ASD</i>	152
5.3.2 <i>Attentional disengagement difficulties in infants at high risk for ASD</i>	153
5.4 <i>Method</i>	156
5.4.1 <i>Gap-overlap task: 7 and 13 months</i>	156
5.4.2 <i>Analysis and measures</i>	157
5.5 <i>Results</i>	159
5.5.1 <i>Prediction of clinical outcome by 13 month measures</i>	159
5.5.2 <i>Relationship with IQ</i>	162
5.6 <i>Discussion</i>	163
<i>Part 2: Interactions between Gaze Following and Disengagement at 7 and 13 months in Infants at High Risk for Autism Spectrum Disorder</i>	166
5.7 <i>Introduction</i>	166
5.7.1 <i>Postnatal brain development</i>	166
5.7.2 <i>Links between disengagement of visual attention and gaze following</i>	169
5.8 <i>Results</i>	171
5.8.1 <i>Relationship between visual disengagement and gaze following measures at 7 and 13 months</i>	171
5.9 <i>Discussion</i>	174
5.10 <i>Conclusion</i>	178
<i>Summary of Chapter 5</i>	179

Chapter 6. General Discussion: Multiple Risk Factors in Autism Spectrum Disorder	180
6.1 <i>Introduction</i>	180
6.2 <i>Aims</i>	180
6.3 <i>Implications: emergence of behavioural symptoms</i>	184
6.4 <i>Alternative Explanations</i>	190
6.5 <i>Limitations and future work</i>	192
6.6 <i>Specific study modifications</i>	196
6.7 <i>Implications for intervention</i>	200
6.8 <i>Conclusion</i>	202

List of Tables

<i>Table 2.1</i> Number of Infants in each Diagnostic Category.....	54
<i>Table 2.2</i> Descriptive Statistics for the Mullen Scales of Early Learning (<i>MSEL</i>) Subscale Standard T-scores and Early Learning Composite Standard Scores (ELC).....	68
<i>Table 3.1</i> Descriptive Statistics for the Mullen Scales of Early Learning (<i>MSEL</i>) Early Learning Composite Standard Scores and ADOS-G Scores.....	89
<i>Table 3.2</i> Number of Valid Trials and Total Duration of Looking (ms) at the 7 and 13 Month Visits by Group and Condition.....	94
<i>Table 3.3</i> Proportion of First Looks to and Attentional Engagement with the Congruent and Incongruent Objects at the 7 Month Visit.....	96
<i>Table 3.4</i> Proportion of First Looks to and Attentional Engagement with the Congruent and Incongruent Objects at the 13 Month Visit.....	97
<i>Table 3.5</i> Correlations between Gaze Following Performance and 24 Month Receptive Vocabulary Size.....	110
<i>Table 4.1</i> Descriptive Statistics for the Mullen Scales of Early Learning (<i>MSEL</i>) and MacArthur Bates Communicative Development Inventory (CDI) Receptive Vocabulary....	127
<i>Table 4.2</i> Correlations between Experimental Performance and Vocabulary Size.....	133
<i>Table 4.3</i> Frequencies for categorical word learning variables.....	140
<i>Table 5.1</i> Descriptive Statistics for Disengagement Scores at 7 and 13 months.....	158
<i>Table 5.2</i> Thirteen Month Looking Time * Disengagement Quartiles Cross-tabulation Split by Outcome Group.....	161
<i>Table A3.1</i> Classification Table for the Four Class Growth Mixture Model.....	210

Figures

<i>Figure 1.1</i> Relationships between behavioural measures over time, and their association with risk status and ASD outcome.	46
<i>Figure 2.1</i> Notation for structural equation models.	63
<i>Figure 2.2</i> Regression of 12 month motor score on 6 month motor score.	65
<i>Figure 2.3</i> The Mplus framework. In this diagram, the circles represent latent variables and the rectangles observed variables. ‘f’ stands for factor, i.e., a continuous latent variable and ‘c’ stands for class, a categorical latent variable. Similarly ‘y’ and ‘u’ represent continuous and categorical observed variables respectively. The arrows specify the different possible relationships that can be modelled between these variables, and the within and between boxes reflect the fact that multilevel models can be run in this framework.	69
<i>Figure 2.4a</i> First order autoregressive model for raw expressive language (EL) score at 7 – 36 month visits.	70
<i>Figure 2.4b</i> Second order autoregressive model for raw EL score at 7 – 36 month visits.	71
<i>Figure 2.4c</i> Latent variable first order autoregressive model for raw EL score at 7 – 36 month visits.	73
<i>Figure 2.5</i> Example of a GCM using EL score at 7 – 36 months.	75
<i>Figure 2.6</i> Graph of the sample (●) and model estimated (▲) means. Because the slope regression coefficients were set at 0, 1, 3, 5, these values on the x-axis correspond to the 7, 13, 24 and 36 month visits respectively.	76
<i>Figure 2.7</i> Example of a GCM using chronological age for EL score at 7 – 36 months.	78
<i>Figure 2.8</i> Local versus global maxima.	80
<i>Figure 3.1</i> Screen shots of the videos presented to infants, split into the three phases used in analysis: looking down, direct gaze, shift; and with the congruent and incongruent object areas of interest (AOIs) highlighted. The visual angle of the overall screen took up 37.6°	

horizontally and 30.5° vertically. Depending on their size, the visual angle of the objects, subtended 3.7° by 4.5° for the smallest and 7.3° by 8.4° for the largest.....	92
<i>Figure 3.2</i> Attentional engagement: Proportion of looking time to the congruent object at the 13 month visit. Error bars +/- 1 S.E.....	98
<i>Figure 3.3</i> Graph showing the relationship between 13 month looking time to the congruent object and CDI receptive vocabulary score at 24 months, with the outlier highlighted.....	111
<i>Figure 4.1</i> Experimental paradigm. Example of objects and words used in fast-mapping and retention trials. * indicates performance significantly above chance at the $p < 0.05$ level.	125
<i>Figure 4.2</i> Performance during the fast-mapping trials. Chance level is 0.33. Error bars +/- 1 standard error. * indicates performance significantly above chance at the $p < 0.05$ level.	129
<i>Figure 4.3</i> Performance during the retention trials. Chance level is 0.50. Error bars +/- 1 standard error. * indicates performance significantly above chance at the $p < 0.05$ level.	130
<i>Figure 4.4</i> Performance during the retention trials, depending on whether the feedback was reinforcing or correcting children's initial choices. Chance level is 0.50. Error bars +/- 1 standard error. * indicates performance significantly above chance at the $p < 0.05$ level.	132
<i>Figure 4.5</i> Path analysis model looking at the relationship between 13 month looking time and 24 month fast mapping and feedback performance, controlling for risk status.....	141
<i>Figure 4.6</i> Path diagram of the relationship between 13 month looking time behaviour and mutual exclusivity and later word learning following reinforcement and correction.....	142
<i>Figure 5.1</i> From Elsabbagh and Johnson (2010). Three different models for how risk factors might lead to ASD.....	150
<i>Figure 5.2</i> A 13-month-old infant watching an overlap trial in the gap-overlap task.....	157

<i>Figure 5.3</i> Growth curve model to look at the relationship between 13 month looking time and disengagement and the <i>MSEL</i> composite score.....	163
<i>Figure 5.4</i> Autoregressive cross-lagged model. Correlations among 7 month gaze following measures and among 13 month gaze following measures are not shown in the diagram. The only significant correlation is between 13 month looking time and 13 month first look ($r = -0.24$, $p = 0.04$).....	171
<i>Figure 5.5a</i> Scatterplot showing the relationship between disengagement and looking time to the congruent object at 7 months.....	172
<i>Figure 5.5b</i> Scatterplot showing the relationship between disengagement and first look difference score at 7 months.....	173
<i>Figure A3.1</i> Four class growth mixture model for <i>MSEL</i> expressive language (EL) scores.....	211
<i>Figure A5.1</i> : Histograms to show the distribution of First Look Difference Scores at 7 and 13 months.....	212
<i>Figure A5.2</i> : Histograms to show the distribution of Looking Time Scores at 7 and 13 months.....	212
<i>Figure A5.3</i> : Histograms to show the distribution of Disengagement Scores at 7 and 13 months.....	212

List of Acronyms

ADHD	– Attention Deficit Hyperactivity Disorder
ADI-R	– Autism Diagnostic Interview-Revised
ADOS-G	– Autism Diagnostic Observation Schedule-Generic
ANOVA	– Analysis of variance
AOI	– Area of interest
AOSI	– Autism Observation Scale for Infants
ASD	– Autism Spectrum Disorder
ASD-sibs	– Autism Spectrum Disorder siblings
AT-sibs	– Atypically Developing siblings
BAP	– Broader Autism Phenotype
BASIS	– British Autism Study of Infant Siblings
BIC	– Bayesian Information Criterion
CDI	– MacArthur-Bates Communicative Development Inventory
CEFT	– Children’s Embedded Figures Test
CFI	– Comparative Fit Index
CI	– Confidence Interval
DAWBA	– Development and Wellbeing Assessment
EEG	– Electroencephalography
EF	– Executive function
EL	– Expressive Language
ELC	– Early Learning Composite
EPF	– Enhanced perceptual functioning
ERP	– Event related potential
ESCS	– Early Social Communication Scales

FFA – Fusiform face area

FM – Fine Motor

fMRI – Functional magnetic resonance imaging

GCM – Growth curve model

GEE – Generalised estimating equation

GM – Gross Motor

GMM – Growth mixture model

ICC – Intra-class correlation

JA – Joint Attention

MAR – Missing at Random

MCAR – Missing Completely at Random

ME – Mutual Exclusivity

ML – Maximum Likelihood

MSEL – Mullen Scales of Early Learning

ms – Milliseconds

OCD – Obsessive-compulsive disorder

OR – Odds ratio

PDD – Pervasive Developmental Disorder

RL – Receptive Language

RMSEA – Root Mean Square Error of Approximation

RRBIs – Restricted and repetitive behaviours

RT – Reaction time

SCQ – Social Communication Questionnaire

SD – Standard deviation

SE – Standard error

SEM – Structural equation modelling

SPSS – Statistical Package for the Social Sciences

STS – Superior temporal sulcus

T1 – Time 1

T2 – Time 2

T3 – Time 3

T4 – Time 4

TD-sibs – Typically Developing siblings

ToM – Theory of Mind

VABS – Vineland Adaptive Behaviour Scales

VR – Visual Reception

WASI – Wechsler Abbreviated Scales of Intelligence

WCC – Weak central coherence

WLSMV – Mean and variance adjusted weighted least squares

Chapter 1

Introduction: Infant Predictors of Later Outcomes in Typical and Atypical Development.

1.1 Introduction

Autism Spectrum Disorder (ASD) is a highly heritable, pervasive developmental disorder characterised by a core set of behavioural impairments in social interaction, communication and patterns of restricted or repetitive behaviours (DSM-IV-TR, American Psychiatric Association, 2000). At present, an ASD diagnosis rarely occurs before 2 years of age. A longitudinal prospective design, following infants at high risk for developing ASD (who have an older sibling with a diagnosis), thus offers an opportunity for understanding the emergence of ASD over development. Being able to characterise and quantify the developmental processes underlying ASD is not only of inherent basic scientific value, but can also inform targeted intervention strategies, before the onset of core diagnostic symptoms.

The primary aim of this thesis is to look at how early social and non-social behaviours in infancy relate to later developmental and clinical outcomes both in typical development and in infants at high risk for an ASD. One of the earliest behavioural signs of ASD is problems sharing attention with another person about an object (joint attention; JA). However, little is known about the precursors to this behaviour and whether, in ASD, difficulties are driven primarily by problems with social understanding (that other people have their own minds, beliefs and interests) or by problems in flexibly switching attention (i.e., disengaging attention from another person's face to look at the object). There is a substantial body of research into JA and visual disengagement both in typical development and in ASD. These behaviours are thus appropriate targets for examination in infants at high risk for ASD, and allow hypotheses motivated by developmental theory to be tested, which is particularly

important when applying structural equation modelling (SEM) techniques (discussed in depth in Chapter 2).

The broad aim of this thesis is to apply a variety of statistical modelling approaches to longitudinal developmental data, from infants at high risk for ASD. The multivariate SEM approach enables the relationships between multiple variables to be modelled simultaneously. Applying this technique thus offers a conceptual step forward from previous studies, which have tended to map only individual risk markers onto ASD outcomes. These statistical techniques also address some of the challenges associated with the analysis of longitudinal data, such as missing data and measurement error. Here, I apply different models to 1) examine the relationship between behavioural markers and clinical outcome at 3 years of age; 2) link early social behaviours with later language outcomes; and 3) characterise developmental trajectories across the first three years of life.

In the current chapter, I will first briefly review the different study designs (e.g., retrospective and prospective) that have been used for linking infancy measures with later outcomes. I will then discuss evidence for infant factors predicting later outcomes in both typical and atypical development. Next, I will review various different theories of ASD, and discuss the predictions these theories might make about the earliest behavioural signs and underlying developmental processes in ASD. The current evidence for early ASD risk markers from longitudinal prospective studies will be considered with respect to these theories of ASD. Finally, I will discuss the specific relationships between early behaviours and later outcomes which are analysed in this thesis.

1.2 Approaches for predicting later outcomes

There are various methodologies that can be used to investigate the influence of early indicators on later outcomes. These methods fall into two categories: retrospective studies,

which involve looking back at early development, and prospective designs, in which children are followed longitudinally. In the case of atypical development, such as ASD, retrospective studies are in many ways easier to run, because the sample group already have a confirmed ASD diagnosis, and because often only a single visit or even phone call is required. However, there are several methodological drawbacks to retrospective studies which are discussed below. While prospective studies are expensive and time-consuming to carry out, they offer various important advantages over the retrospective method.

One of the most commonly used methods to look for early predictors - particularly of atypical development - is parental report, as the information is relatively easy to collect. However, parent report measures are subject to several methodological drawbacks, for example, they are sensitive to parents' knowledge about typical and atypical development, and the retrospective nature of these reports makes them subject to recollection bias. For example, the tendency to report past events as having occurred more recently than they did, 'telescoping', is particularly problematic when considering retrospective data on developmental milestones (Lord, Risi & Pickles, 2004). Further, it is also plausible that, given a particular diagnostic outcome, parents may misattribute the existence of current symptoms to earlier in development, or alternatively, if parents believe that their infant's development regressed, they may report no early symptoms at all, in line with this view.

Another type of retrospective design involves the analysis of home videotapes. This method has some advantages over parental report, namely that researchers can look for particular behaviours in a more controlled, experimental manner. However, the videos themselves are likely to be variable – children are not filmed randomly but at particular events such as birthdays, and thus videos may not reflect the child's typical behaviour.

One type of prospective approach is a population study, which uses screening tools to identify those at increased risk for having a particular disorder. This method is useful for

looking at particular characteristics without introducing retrospective bias, but population studies require large sample sizes and limited resources mean that only certain methods can be employed (for example, brain imaging would not be a feasible option).

Another approach, and the one used to collect the data in this thesis, is to run a prospective longitudinal design following infants at increased familial risk for developing a particular disorder (owing to having an older sibling or a parent who has the disorder). For prospective studies of infants at high risk for ASD, this is typically an older sibling with an ASD diagnosis. The risk for ASD in siblings, whilst low in absolute terms, is still much higher than in the general population (~5–10%; Bolton et al., 1994; Constantino, Zhang, Frazier, Abbacchi & Law, 2010). A recent paper looking at estimates in high-risk sibling studies found recurrence rates of just below 20% (Ozonoff et al., 2011). Whilst potential biases from the sample were accounted for in this study, it is important to note that ASD outcome was measured at 3 years of age, and it is possible that stability of these diagnoses will shift over the course of development. It will be important to follow up these children and see whether similar recurrence rates persist into middle childhood. However, while there may be debate over the precise estimates, risk is certainly elevated in siblings, and the prospective high-risk paradigm provides the opportunity to collect extensive data using a range of techniques and methodologies to compare the development of high-risk infants (who have an older sibling with an ASD) and low-risk controls (with a typically developing older sibling).

As well as the high-risk children who go on to develop ASD, there are also a large proportion of the children who show subclinical symptoms of the disorder. These might be difficulties restricted to a single domain (i.e., only restricted and repetitive behaviours), or sub-threshold symptoms across domains. This group of children has sometimes been referred to in the literature as the broader autism phenotype (BAP). This BAP concept comes from evidence that genetic risk for ASD can result in brain and behaviour characteristics associated

with the autism phenotype not only in affected individuals but also, to some degree, in genetic relatives. Bolton et al. (1994) showed that, using a family history interview, first-degree relatives of autistic individuals were more likely to show social-communication problems and stereotyped behaviours than those with a relative with Down syndrome. There have been attempts to define the BAP in adulthood, both as a clinical category (e.g., Losh & Piven, 2007) and as characteristics extending to the general population (Constantino & Todd, 2003). However, there is no agreed definition about what exactly constitutes the BAP in infancy and early childhood (Rogers, 2009), and the term can also refer to group level differences between low-risk and high-risk infants. Many risk group differences have been found in socio-communicative and cognitive behaviours (see Elsabbagh & Johnson, 2007 for a review). It is of course possible that these differences are being driven by a few extreme individuals (i.e., potentially from those of infants who later go on to develop ASD). However, the distribution of data (e.g., Elsabbagh et al., 2009a; Presmanes, Walden, Stone and Yoder, 2007) suggests that the group effect is not necessarily being driven by outliers. Thus risk status could be referred to as a BAP effect.

In this thesis, when referring to the findings of other studies, any use of the term BAP will be qualified. However, given its mixed usage, the approach taken here is to 1) compare high-risk and low-risk groups to look for a ‘group effect’; and 2) compare clinical outcome groups (explained in detail in Chapter 2) with the high-risk group split into those with ASD, those with atypical development and those who appear typically developing at 3 years of age.

1.3 Infant predictors of later outcomes in typical development

The search for early infant predictors of subsequent outcomes, both in typical and atypical development, has been an area of research interest for over 50 years. Outcome measures have included language abilities and intelligence test scores, as well as categorical measures of

atypical development (i.e., disorder versus no disorder). The prediction of measures over time can inform developmental theory and further understanding of the potential underlying mechanisms in typical neural and cognitive development. Further, the ability to predict these later outcomes has implications in terms of the appropriate and early targeting of potential intervention strategies.

One of the most commonly examined outcome measures is intelligence score. Before 1960, intelligence was assumed to be a stable construct from infancy into adulthood. However, Hunt (1961) suggested that intelligence, as measured by standardised intelligence tests, may in fact change across the lifespan. In a review of the literature, Bornstein and Sigman (1986) showed that the average correlation between scores on standardised tests of development in infants aged 1–6 months, and later intelligence at 5–7 years, was very low ($r = 0.09$). However, while standardised tests within the first year of life did not reliably predict later intelligence, the search for possible ‘early markers’ continued. McCall and Carriger (1993) in a meta-analysis of the literature found that infant performance in both habituation and recognition paradigms predicted subsequent intelligence scores.

Habituation tasks involve repeated viewing of a ‘standard’ stimulus until there is a reduction in infant looking time (typically to 50% of that for initial presentation of the stimuli), followed by the simultaneous presentation of the standard stimulus and a novel stimulus. Recognition memory paradigms are similar but involve the presentation of pairs of identical stimuli (the ‘standard’), followed by presentation of a pair of stimuli containing the standard stimulus and a novel stimulus. These types of task are thought to reflect information processing skills such as speed and accuracy as well as stimulus encoding abilities (McCall & Carriger, 1993).

One of the first tasks designed to link individual differences in early cognitive development with subsequent performance on intelligence tests was the Fagan Test of Infant

Intelligence. The rationale behind this test was the belief that intelligence consists of innate, automatic processes which enable the acquisition of knowledge. Through interaction with the environment, over the course of development, individual differences in the speed of early processing will predict knowledge acquisition as measured by later intelligence tests (see Fagan, 1984). The main part of the Fagan test uses a habituation paradigm, termed “novelty problems”. Following habituation, infant looking time to the standard and novel stimuli is measured, and a novelty score calculated. Fagan and Montie (1988) showed a moderate correlation ($r = 0.49$) between performance on visual novelty problems between 3–7 months of age and subsequent score on a standardised intelligence test at 3 years of age.

Understanding the underlying developmental mechanisms is important when thinking about why some measures reliably predict later IQ whereas others do not. The fact that infant standardised tests predict so little of the later variance in IQ could be either because they are not measuring the same underlying ability, or because IQ is not a stable construct over the first few years of life. These are very different interpretations of the same findings. Thus, it is necessary to have a clear hypothesis about the mechanisms involved and whether an early behaviour is thought to be a prerequisite versus an earlier manifestation of the same behaviour. McCall (1994) argues that it is an infant’s ability to disengage attention that is mediating the relationship between habituation and IQ, rather than there being a direct relationship per se.

As well as looking at links with later intelligence test scores, other more recent studies have mapped early behavioural characteristics, such as joint attention (JA) and disengagement of visual attention, onto a range of more specific outcomes, including language and temperament. In this case the aim is not to measure the continuity of the same construct across development, but to examine early behaviours as prerequisites for later skills.

Some links between early measures of visual attention and cognitive factors have been reported, although the majority of these are cross-sectional findings. For example, Johnson, Posner and Rothbart (1991) found that infants who were better able to disengage attention were more 'soothable'. Further, McConnell and Bryson (2005) demonstrated that increased latency to disengage at 6 months was associated with greater levels of frustration at the same age. However, there were no longitudinal links found for earlier measures of disengagement (at 2 or 4 months of age). Courage, Howe and Squires (2004) showed longitudinal links between visual attention at 3.5 months and recognition memory at 1 year of age. They found that shorter look duration, which could be an index of faster disengagement, was related to enhanced recognition memory performance.

Perhaps the most widely studied predictor of later language abilities is early JA behaviour. JA is defined as shared engagement with another person (Mundy, Sigman, Ungerer & Sherman, 1986) and requires the ability to switch attention between a person and a referred object. Being able to engage in this type of shared attention is likely to be important in developing word-object associations, and in this sense JA can be seen as a prerequisite ability for later language development. However, JA itself is also a measure social-communication skill, and thus could represent an early manifestation of communication ability. In this latter sense, JA could be seen as a precursor to later language ability, with both skills reflecting the same underlying ability over time. Evidence for a link between early JA and later language is discussed more fully in Chapter 3.

Examining the relationship between measures over time furthers our understanding of the processes and mechanisms underlying typical development. Studying the emergence of behaviours in atypical development, as well as extending what is known about specific disorders, can also inform our understanding of typical developmental processes, and allow specific hypotheses about the underlying mechanisms to be tested.

1.4 Predictors of later outcomes in atypical development

Studies using high-risk samples capitalise on the genetic component of many disorders to increase the sample size of participants who later develop the disorder. Longitudinally following these high-risk children, who have a parent or sibling with a particular disorder, enables early predictors of the disorder to be identified, as well as detailed characterisation of the premorbid phase, before the onset of overt symptoms.

For example, children of a parent with bipolar disorder are at increased risk of developing the disorder themselves as well as other forms of psychopathology, and the use of longitudinal studies has begun to elucidate some of the precursors to bipolar disorder (Chang, Steiner & Ketter, 2000). Children of parents with schizophrenia are also at elevated risk for psychopathology and studies with these high-risk children have also found delays in motor skills and verbal memory, and difficulties in smooth-pursuit eye movements (Erlenmeyer-Kimling, 2000; Schubert & McNeil, 2004).

Other high-risk studies have looked at disorders more commonly associated with a childhood diagnosis, including dyslexia and attention deficit hyperactivity disorder (ADHD). This high-risk strategy is of particular value when studying disorders that typically emerge in early childhood, because characterising the developmental trajectories leading to symptoms may allow for intervention while there is still a great degree of plasticity in neural development. In comparison to low-risk controls, siblings of children with ADHD have been found to show clinically increased behavioural, mood and anxiety disorders (Faraone, Biederman, Mennin, Gershon & Tsuang, 1996). In children at high risk for dyslexia who have a dyslexic parent or sibling, subclinical problems have been found, with children showing difficulty with phonological processing despite typical language development (Carroll & Snowling, 2004).

Taken together, studies predicting outcomes in typical and atypical development further our understanding of the processes and mechanisms that might underlie the development of particular behaviours, as well as how developmental trajectories can be shifted by a combination of genetic and environmental risk factors. These results generate hypotheses about potential candidate mechanisms that could underlie symptomatology in ASD, and allow a targeted, theory-driven approach to be applied when developing research questions.

1.5 Predictors of the development of ASD

The question of when ASD emerges has been of interest to researchers since Kanner's (1943) description of the disorder, in which he argued that autism is present from the first few months of life. Understanding how ASD manifests in infancy is important both for predicting later outcomes (such as IQ or social functioning) as well as for developing early interventions, which can take advantage of the fact that the brain is highly plastic during infancy. Developmental models emphasise sequential and mutually interacting processes by which biological background and environmental input continuously interact to shape cognitive development and behaviour, which in turn can change gene expression and the environmental input. Thus the influence of an intervention early in development has the potential to be magnified in a positive feedback loop.

For many years the only information about early signs of ASD was provided by retrospective parent reports, which suggested that JA behaviours are among the best early 'discriminators' of ASD, in infants between 12 and 18 months (Charman, 2003). Converging evidence for this came from another retrospective design, home videotape analysis from children's first birthday parties by Osterling and Dawson (1994). They found that, compared with typically developing children, those who went on to develop ASD were less likely to

look at other people, orient to their name and show or point to objects – skills all related to JA.

Other studies have used a prospective design with instruments devised to screen for ASD in the general population. Baird et al. (2000), using the CHAT screening questionnaire in a population study of 18-month-old infants, found that, together with a lack of imaginative play, gaze monitoring and pointing for interest (again joint attention-type behaviours) predicted ASD outcome. Measurement sensitivity is important to note here, as only 38% of the children who went on to develop ASD were identified using this method. This implies that, whilst deficits in JA are highly predictive of ASD, typical JA does not preclude a diagnosis of ASD. However, 18 months is relatively late to be measuring JA, because in typical development these behaviours begin to emerge towards the end of the first year of life (e.g., Scaife & Bruner, 1975; Tomasello, Carpenter, Call, Behne & Moll, 2005).

Another methodology, which avoids many of the discussed limitations, is the high-risk sibling prospective design. In order to explore the development throughout infancy of a range of brain and behaviour characteristics related to ASD, prospective longitudinal studies of high-risk infants are required. There are now many published studies which have used the high-risk paradigm to look at group level differences. However, far fewer studies have been published that look at early predictors of later ASD outcomes. The emerging findings suggest that overt behavioural symptoms are extremely subtle within the first year of life. From 12 months of age, however, a range of behavioural impairments in both social and non-social domains have been observed and shown to relate to the later development of ASD, and these are discussed below. Several studies (e.g., Zwaigenbaum et al., 2005; Ozonoff et al., 2010) have also demonstrated that change over time from around 6–12 months is predictive of later symptoms of ASD, suggesting that the trajectory of development may also be important.

The scope of the current thesis is the prediction of later outcomes. Therefore, I will focus on those prospective studies which have linked early social and non-social measures to later clinical ASD outcome at 3 years of age. In the following section I will discuss some of the key theories of ASD and consider, in light of these theories, the emerging findings from prospective studies of infants at high risk for ASD.

1.6 Social development

Perhaps one of the most striking features of children with an ASD is their difficulty with understanding the subtleties of social interaction. There are various theories, both biological and cognitive, which suggest the primary deficit in ASD is a social one. In typical development, early social interaction (e.g., attention to faces, reciprocal smiling) is thought to be important from an evolutionary point of view for bonding with the caregiver in order to be looked after (Morton & Johnson, 1991) as well as for learning. Compared to other species, new-born human infants are particularly vulnerable with a protracted period of post-natal development that necessitates round the clock care, but this also offers human infants an advantage over more independent ones, because it is this prolonged dependence on caregivers that also increases the chances to learn from them. Csibra and Gergely (2006) have argued that learning from other people is potentially the most important human adaptation.

Perhaps one of the most influential cognitive theories of ASD over the previous two decades has been Theory of Mind (ToM). ToM refers to an understanding that other people have their own minds, i.e., desires, beliefs and their own interpretations of the world, and is argued by some (e.g., Baron-Cohen, Leslie & Frith, 1985) to be impaired in individuals with ASD. The argument, broadly, is that difficulties in representing the beliefs of others may lead to difficulty predicting other people's behaviour, and potentially a withdrawal from the social world. ToM has typically been tested in young children using false belief paradigms

involving puppets. In Baron-Cohen et al.'s (1985) classic Sally–Anne task, two dolls, Sally and Anne, are together in a room and Sally puts a marble in a basket, and then leaves the room. While she is outside the room, Anne moves the marble into a box. Sally then returns and the children are first asked two control questions ‘Where is the marble really?’ and ‘Where was the marble at the beginning?’ to check their understanding and memory for the story. They are then asked the key false belief question ‘Where will Sally look for her marble?’ If children are able to represent that Sally has a belief (the marble is in the basket where she left it) that is different to reality (the marble is actually in the box), then they should answer the false-belief question correctly. Baron-Cohen et al. (1985) showed that compared to 85% of typically developing children, and children with Down syndrome, only 20% of children with ASD pass the false-belief task.

There are various criticisms of the idea that ToM is the core impairment in ASD. One problem relates to specificity of the ToM account to ASD, as even in Baron-Cohen et al.'s (1985) original study, 20% of children with an ASD still pass the false-belief task. It is possible that the understanding of ToM is actually a more graded process and individuals with ASD show a degree of difficulty rather than an overall impairment. Baron-Cohen, Jolliffe, Mortimore and Robertson (1997) demonstrated that even high functioning adults with ASD show problems in inferring mental states from photographs of eyes. However, another issue relates to the debate over when ToM emerges, and what factors drive its development. For ToM to be an etiological factor in the development of ASD, impairments must be present from very early on in development. However, passing classic measures of ToM, such as the Sally–Anne task does not occur until 3–4 years of age even in typical development. A more recent study used eye-tracking to record children's spontaneous gaze behaviour and found evidence for false belief understanding much earlier, at 2 years of age

(Southgate, Senju & Csibra, 2007). However, this is still much older than when the earliest behavioural signs of ASD are being identified.

Perhaps then ToM itself is the result of difficulties with earlier behaviours. For example, if infants with ASD do not orient towards another person's face, they will miss out on a range of cues likely to be related to ToM, such as a person's emotional state and the focus of their attention. A longitudinal study in typically developing children by Charman et al. (2000) showed links between JA behaviours (gaze switching between an adult and a toy) at 20 months of age, and subsequent ToM at 44 months. Baron-Cohen (1989a) argued that JA behaviours reflect the same underlying processes involved in ToM, and JA can be conceptualised as an early manifestation of metarepresentation, rather than a precursor to it.

A final problem with ToM as a causal account is that, while it may be able to explain some of the social-communication problems in ASD, there is little evidence for ToM playing a role in restricted and repetitive behaviours (RRBIs). In a review of the literature Happé and Ronald (2008) note that while social withdrawal may play a role in developing unusual topics of interest, there is little evidence that such behaviours result from a lack of social understanding, as no relationship between social functioning and number of restricted and repetitive behaviours (RRBIs) has been found. And further, repetitive behaviours in ASD are thought to be rewarding in themselves, with children often displaying such behaviours when they are alone, not just in social situations.

Chevallier, Kohls, Troiani, Brodtkin and Schultz (2012) argue that ASD is better conceptualised as a reduction in social motivation. In their theory, social motivation at the behavioural level is comprised of social orienting (e.g., attention to faces, preference for direct gaze), seeking-liking (e.g., values of social reward) and social maintaining (e.g., reputation management). The authors argue that the social-motivation theory, unlike ToM, is a causal account, in which reduced interest in the social world results in fewer social inputs

during development leading eventually to a reduced level of social-cognitive skill. Under this account, ToM difficulties arise as a secondary consequence of a reduced interest in the social world. However, it seems to me that it is equally plausible, at least from a theoretical perspective, that problems in understanding other people could lead to a reduction in social motivation.

The mechanism proposed to underlie the behavioural level of social motivation theory is a problem with the functioning of the orbitofrontal-striatum-amygdala network, which leads to difficulties in representing the reward value of social stimuli (e.g., Dichter, Richey, Rittenberg, Sabatino & Bodfish, 2012; Scott-Van Zeeland, 2010). However, whether this reward processing deficit is specific to the social domain is a little unclear because the control task typically used relates to gambling or monetary reward, which arguably exists within a social context.

Like social motivation theory, the ‘social brain hypothesis’ also emphasises the neural underpinnings of social behaviour, suggesting that there is a dedicated neural network subserving the processing of social information. This network is thought to include the fusiform face area (FFA), superior temporal sulcus (STS), and regions of orbitofrontal cortex (e.g., Adolphs, 2003), with regions specialised for particular functions but working together to process social stimuli. Hadjikhani, Joseph, Snyder and Tager-Flusberg (2007) demonstrated that adults with ASD show activation in some regions within the social brain network, such as the FFA when viewing faces. However, the pattern of activation across regions in the overall network showed atypical activation.

While event related potential (ERP) methods are not well suited to address the question of where particular functions are localised within the brain, they are useful for looking at brain responses in infancy and provide converging evidence for the idea of developing brain specialisation. In adults there is a face-sensitive ERP component, the N170, which peaks

around 170 milliseconds (ms) after the onset of a stimulus. A similar, precursor component has been found in infants, the N290, although this component has a longer peak latency and a smaller amplitude than the adult N170 (de Hann, Pascalis & Johnson, 2002; Halit, Csibra, Volein & Johnson, 2004). ERP studies have looked at the modulation of this face-sensitive N170/N290 in response to eye-gaze direction. The component is modulated by the direction of eye-gaze in typically developing 4-month-olds, but not in young children or adults (Johnson et al., 2005). Modulation of the N170/N290 by the direction of eye-gaze suggests that the same neural regions underlie both eye-gaze and face processing. However, Grice et al. (2005) demonstrated that young children (aged 2–5 years) with ASD still show modulation of the N170 in response to direction of eye-gaze. This could suggest that the specialisation process, which leads to somewhat separable processing of faces and eye-gaze direction by typical adults, may not be occurring in ASD.

This is an attractive theory, although we are still faced with the question of why specialisation for social functions does not occur to the same extent in ASD. It is difficult to know the direction of causality; whether a lack of social interest results in delayed or reduced neural specialisation, or alternatively a lack of specialisation results in reduced processing of social information.

Other developmental models of ASD suggest that the primary deficit is an early difficulty with social processing (e.g., Mundy & Neal, 2001) or social attention, such as JA and social orienting (e.g., Dawson, Meltzoff, Osterling, Rinaldi & Brown, 1998) leading to reduced social feedback and increasing social problems. Both theories are based on the idea that early social difficulties are likely to lead to a ‘negative feedback loop’ in which social problems become compounded over the course of development.

While the above theories approach the development of ASD at different levels (i.e., behavioural, cognitive and neural) they have in common their attempt to explain the disorder

in terms of a primarily ‘social’ deficit. Fundamentally, they argue that there is a difference in the processing of social stimuli in ASD. This may arise owing to a lack of understanding of other people’s minds, a lack of reward or motivation relating to social stimuli or a lack of specialisation of the social brain.

1.6.1 Prospective social studies

While it is clear that social problems characterise children with ASD, and social deficits have been proposed as critical in various theories of ASD, less is known about the early development of social behaviours in infants at high risk for an ASD. Understanding the emergence of social behaviours can inform and reshape existing theories. This brief review of the literature will focus on the findings from prospective studies of high-risk infants which look at the relationship between early social behaviours and later ASD outcome.

JA type behaviours were among the first to be investigated by prospective high-risk studies, as there is evidence that JA is one of the earliest signs of ASD (e.g., Charman, 2003; Osterling and Dawson, 1994). Sullivan et al. (2007) showed response to JA at 14 months to be a significant predictor of ASD outcome. Yoder, Stone, Walden and Malesa (2009) also found a relationship between 15 month responding to JA and later ASD outcome. Further, the growth rate of a weighted triadic communication measure derived by multiplying the frequency of communication bids by the behaviour used (i.e., 1 = nonverbal, 2 = single word, 3 = multi-word) was also associated with ASD diagnosis.

This relationship between JA and ASD outcome is consistent with the ‘social’ theories of ASD. According to extensions of ToM, JA reflects the same underlying construct – metarepresentational understanding – and should thus be impaired early in development. The ‘social brain’ hypothesis suggests there is reduced specialisation of the networks underlying social processing, and thus social behaviours, such as JA, should be impaired. Social orienting theories specifically predict difficulties in social attention, including JA. And

finally, social motivation theory also predicts JA difficulties, although in this case JA problems are classified as a secondary deficit, resulting as a consequence of reduced motivation.

In a more recent study Elsabbagh et al. (2012) looked at response to another person's eye-gaze in infancy, a potential precursor to later JA problems. They used ERPs to examine neural correlates of response to dynamic gaze shifts either toward or away from the infant. A lack of discrimination between the two conditions (direct versus averted), in terms of ERP components at 6–10 months distinguished those infants who later went on to develop ASD from low-risk controls and other high-risk infants. Importantly this finding suggests that it is possible to differentiate those infants who later go on to an ASD diagnosis *before* the onset of overt behavioural symptoms. The authors argue that the dynamic stimuli used in this study are likely to recruit multiple brain regions from the social brain network. The lack of differentiation in the group who later develop ASD could reflect reduced specialisation of the social brain.

Nadig et al. (2007) found that a failure to respond to name, a measure of social orienting, in 1-year-olds showed specificity (i.e., the percentage of children without ASD correctly identified as not having the disorder) of 89% for ASD outcomes at 24 months and 94% for all types of developmental delay. However, sensitivity (i.e., the percentage of children with ASD correctly identified as having the disorder) was not as high, at 50% for ASD and 39% for developmental delay. These findings suggest that there is some degree of prediction of symptomatology from measures of social orienting, supporting the idea of an early difficulty with social attention.

Landa, Holman and Garrett-Mayer (2007) found differences in social, communication and play behaviours at 14 months in children showing symptoms of ASD at 14 months (an early diagnosis group), in comparison to all other groups (i.e., low-risk, ASD later-diagnosis at 36

months, high-risk BAP and high-risk non-BAP). In this study, the 'BAP group' showed language or social difficulties but did not meet clinical criteria for ASD, and are thus similar to the atypically developing-siblings in this thesis, described fully in Chapter 2. Another study, Ozonoff et al. (2010) looked at a range of early social communication behaviours, including looking at faces, shared smiling and directed vocalisations. Consistent with results from other social behavioural measures, at 6 months no group differences were found between low-risk controls and those who later developed ASD. However, during the following 12 month period, those in the ASD group showed declining trajectories and loss of skills across social communication measures. Young et al. (2011) also showed delayed imitation at 12–24 months in children who later developed ASD compared to low-risk controls, but not other high-risk infants.

Finally, in terms of temperament, Garon et al. (2009) found that 24-month-old children subsequently diagnosed with ASD showed lower positive affect and higher negative affect than high-risk siblings without an ASD diagnosis and low-risk controls. The ASD group also showed a temperament profile characterised by decreased sensitivity to social reward (reduced Behavioural Approach). Another study (Hutman et al., 2010) looked at infant responsiveness to another's distress and found that children in the ASD group showed reduced change in affect (compared to high-risk and low-risk children) when the experimenter pretended to hit her finger with a toy hammer at 12, 18, 24 and 36 months.

None of these studies have found impairments in social-communication behaviours within the first year of life, although Elsabbagh et al.'s (2012) electroencephalography (EEG) study is the first to find evidence for an atypical brain response to social stimuli, before the onset of overt behavioural differences. The lack of early behavioural differences in the first year of life provides evidence against social motivation theory. If motivation causes less interest in social stimuli this is likely to have a knock-on effect on the brain regions underlying social

processing. However, the opposite pattern, with early lack of neural differentiation in response to social stimuli before behavioural symptoms emerge, is more in line with the predictions of the social brain hypothesis. Elsabbagh et al. (2012) argue that the dynamic stimuli used in this study are likely to recruit multiple brain regions from the social brain network. The lack of differentiation in the group who later develop ASD could reflect reduced specialisation of the social brain.

1.7 Non-social development

ASD is defined by a triad of behaviours and as well as social-communication difficulties, children with ASD also show patterns of restricted and repetitive behaviours. These can include obsessive interests, such as knowing the names of all the dinosaurs, stereotyped and repetitive motor behaviours (e.g., hand flapping), a need for routine and sameness and a detail focused cognitive style. There are several theories which have emphasised non-social behaviour as the central problem in ASD, including executive dysfunction, perceptual theories and underconnectivity theory.

Executive functions (EFs) refers to a range of behaviours such as planning, online monitoring, inhibition and working memory, thought to be mediated primarily by the prefrontal cortex (Kramer & Quitania, 2007; Stuss, 2007). The 'executive dysfunction' account of ASD is thus domain general, unlike ToM for example which proposes a specific underlying deficit. There is evidence that EFs are an area of difficulty in ASD, although deficits are not found universally. A review of the literature (Hill, 2004) suggested that planning difficulties and reduced mental flexibility (i.e., perseverative behaviours) characterise individuals with ASD irrespective of age or ability, but other abilities such as inhibitory control are sometimes unimpaired. This mixed profile, together with the fact that EF difficulties are also characteristic of several disorders including ADHD, obsessive-

compulsive disorder (OCD), Tourette syndrome and schizophrenia, makes it difficult to see how executive dysfunction could be the primary deficit in ASD. Further, Lopez, Lincoln, Ozonoff and Lai (2005) argue that although EFs relate to non-social aspects of ASD, including motor stereotypies, this account cannot explain the full triad of autistic symptoms.

However, Russell, Mauthner, Sharpe and Tidswell (1991) demonstrated links between EFs and ToM, and argued that autistic children's failure in the Sally–Anne task was owing to an inability to inhibit responding to where the marble really is. Russell et al. (1991) devised the 'Windows task' in which children are shown two windows, one with chocolate behind it and the other one empty. In order to win the chocolate, children must learn to inhibit their prepotent response to point at the chocolate, and in fact to point to the empty window. Typically developing 3-year-old children fail this task, repeatedly pointing to the window with chocolate behind it. Russell et al. (1991) claim that this task involves both EFs (inhibiting a direct point to the chocolate) and ToM (the task involves deception because the other player is acting based on the participant's gestures). They suggest that EF is a necessary precursor to ToM, a view supported by Pellicano's (2007) finding that some children with ASD fail ToM and perform well on EF tasks, but not the other way around. However, evidence for consistent deficits in intention monitoring, the proposed mechanism by which EF influences ToM performance, were not found by Russell and Hill (2001).

Another important domain-general cognitive account of non-social difficulties in ASD is weak central coherence (WCC; Frith, 1989; Frith & Happé, 1994). Unlike typically developing individuals who process information in terms of overall meaning by extracting the 'gestalt', according to WCC, ASD is characterised by a cognitive processing style which favours local over global level processing. Evidence for superior local processing in ASD comes from autistic children's above average performance on the Children's Embedded Figures Test (CEFT, Witkin, Oltman, Raskin & Karp, 1971). Shah and Frith (1993) found

that autistic individuals also show better performance, compared to typically developing and mental age matched controls, on the Block Design subtest of the Wechsler Abbreviated Scales of Intelligence (WASI, Wechsler, 1999). In both of these tasks the figure can be split into smaller constituent parts and Frith (2003) suggests that the lack of 'drive' for central coherence in individuals with ASD enables them to outperform typical controls.

WCC was initially proposed to account for ToM problems in ASD, with the suggestion that social information foraging requires high levels of integration and is thus likely to be disproportionately impaired (Frith, 1989). However, Frith and Happé (1994) subsequently noted that WCC was largely independent from success on ToM tasks. Again then, this domain-general account is broadly limited to explaining the non-social symptoms of ASD (Happé & Frith, 2006).

An alternative framework for understanding the perceptual processing abilities seen in ASD was provided by Mottron and Burack (2001) in their Enhanced Perceptual Functioning (EPF) account. While this 'theory' is a fairly comprehensive account detailing the strengths of processing in individuals with ASD, it is largely descriptive, rather than aiming to explain the symptoms of ASD. The EPF account argues that there is enhanced processing in ASD, stating that whilst autistic individuals show increased attention to detail, performance in global and configural processing tasks is typical (Mottron, Burack, Stauder & Robaey 1999). One line of evidence which supports their theory comes from hierarchical figures (letters or numbers which are composed of smaller letters or numbers). Mottron, Burack, Iarocci, Belleville and Enns (2003) found that both autistic and typically developing children showed the same pattern of performance, with a local bias for large letters and a global bias for small letters.

Mottron et al. (2003) also found no group differences on a configural grouping task. This is in contrast to Plaisted, Swettenham and Rees' (1999) divided attention task where more global errors were found in the ASD group. Mottron et al. (2003) suggest that this could be because whilst their grouping task could be resolved using the less consciously accessible dorsal stream, Plaisted et al.'s (1999) task required stimulus identification and thus ventral stream activity (Goodale & Milner, 1992). Bertone, Mottron, Jelenic and Faubert (2005) suggest that rather than ASD being characterised by a dorsal deficit, as suggested by Spencer et al. (2000) based on findings of decreased sensitivity to complex motion in individuals with ASD, a 'neuro-integrative' approach better captures visuo-spatial processing abilities in ASD. Bertone et al. (2005) found that individuals with ASD were better able to discriminate orientation based on first-order rather than second-order gratings. Bertone et al. (2005) argue that this can be explained by atypical connectivity in low-level visual areas, suggesting that increased lateral inhibition could explain the pattern of enhanced performance on first-order gratings and decreased performance on the second-order gratings that require integration across brain areas. Alternatively, according to the 'underconnectivity' hypothesis (Just, Cherkassky, Keller & Minshew, 2004), processing of second-order gratings could be impaired to a greater extent in ASD because complex information requires the combined processing of several different brain regions. In other words, there is reduced efficiency in neuro-integrative mechanisms at higher levels of the visual streams.

Connectivity theories were initially proposed to account for the observed perceptual abilities in ASD. Brock, Brown, Boucher and Rippon's (2002) temporal binding connectivity account was suggested as a potential mechanism underlying the weak central coherence observed in ASD. There is growing evidence for atypical neural connectivity in ASD. However, while some studies have found 'under' connectivity (e.g., Brock et al., 2002; Just et al., 2004), others have found 'over' connectivity (e.g., Rubenstein & Merzenich, 2003,

Belmonte et al., 2004a). According to Belmonte et al. (2004b) this debate has been complicated still further by the different meanings of connectivity, i.e., local versus long-distance cortical connections, as well as connectivity at the synaptic level versus computational connectivity associated with transfer of information. At present, perhaps the most widely agreed upon conception of connectivity theory is that in ASD there is weak connectivity across distal brain regions, but enhanced local connections. Evidence for reduced connectivity and integration across large-scale cortical networks in language processing was provided by Just et al. (2004). Further, the cerebellum is known to play an important role in the organisation of long-range connectivity and both functional and anatomical cerebellar abnormalities are often reported in ASD (e.g., Courchesne, 1997).

Although initially connectivity theory was proposed to account for non-social behaviours, it also offers potential explanations for many of the social difficulties in ASD. For example, in line with the predictions of connectivity theory, the initiation of JA behaviours, which require integration of activity between the posterior and anterior attention systems, are particularly impaired in ASD. Further, Just et al. (2004) proposed that reduced functional connectivity between Broca's and Wernicke's areas in a sentence comprehension task was responsible for language problems in ASD. Further, they hypothesise that ToM problems could result from difficulties in integrating complex social information, although the exact neural substrate of this is somewhat unclear.

1.7.1 Prospective non-social studies

In comparison to the 'social' high-risk prospective studies, there are relatively few looking at non-social behaviours in relation to subsequent ASD diagnostic outcome. Those studies which did choose to focus on non-social abilities have looked at repetitive behaviours, IQ scores, visual attention and play behaviours. While play involving another adult is an

inherently social behaviour, it is included in this section for the repetitive non-social nature of play in ASD.

Christensen et al. (2010) found a reduction in both functional and non-functional play during a free-play session at 18 months in ASD-sibs compared to low-risk controls. More specifically, the ASD-sibs showed less self-directed and other-directed play behaviours and greater non-functional repetitive play behaviours, such as banging or mouthing objects. Comparisons with other high-risk groups were not performed. In another study looking at object exploration, Ozonoff et al. (2008) found that 12-month-olds who were later diagnosed with ASD showed more spinning, rotating and unusual visual exploration of the objects than either the low-risk controls or other high-risk infants. Some of these behaviours, such as unusual visual object exploration, fit in with the WCC framework, which proposes a focus on detail, rather than the whole object. Repetitive behaviours such as banging and rotating objects on the other hand could potentially be explained by a preference for contingency. Klin, Lin, Gorrindo, Ramsay and Jones (2009) found that 2-year-old children with ASD show a preference for stimuli in which the audio-visual information is contingent (i.e., a person clapping, where the motion coincides with the clapping noise).

Loh et al. (2007) coded videos from the Autism Observation Scale for Infants (AOSI; Bryson, Zwaigenbaum, McDermott, Rombough & Brian, 2008) in infants aged 12 and 18 months, and found an increase in stereotyped behaviours in those who later developed ASD. At 12 months ASD-sibs showed greater frequency of arm waving compared to low-risk controls, but not other risk groups. However, by 18 months the ASD-sibs differed from both comparison groups on frequency of arm waving and, both ASD-sibs and other high-risk infants showed increased “hands over ears” postures compared to the low-risk controls.

There is also some evidence for motor problems, as well as other general developmental problems, from studies using the Mullen Scales of Early Learning (*MSEL*; Mullen, 1995), a

standardised developmental assessment. Landa and Garrett-Mayer (2006) found that ASD-sibs performed more poorly than unaffected high-risk children at 14 months on all scales except visual reception (VR) (i.e., gross motor; GM, fine motor; FM, receptive language; RL and expressive language; EL). By 24 months, the ASD-sibs scored lower than unaffected children on all scales, and lower than language delayed children on GM, FM and RL. Finally the ASD-sibs showed a slower developmental trajectory with decreased development from 12–24 months.

Finally, one study has looked at the relationship between early visual attention and subsequent ASD outcome at 36 months (Elsabbagh, Fernandes et al., under review). This study looked at visual disengagement from a centrally presented stimulus to a peripheral one in 7- and 13-month-olds at high risk for ASD. Those infants who later went on to receive an ASD showed an increase in disengagement latencies from the 7–13 month visit, while children in the other groups became faster. Slower disengagement from a stimulus, so-called ‘sticky fixation’, results in prolonged fixation durations. This lack of flexibility when scanning a visual scene could plausibly lead to difficulties with attentional control, as proposed by EF, and a more local, detail-focused cognitive style.

1.8 What comes first: social or non-social?

The theories discussed above tend to fall into the category of ‘social-first’ or ‘non-social first’, with a core deficit in one domain leading to downstream effects in a different domain. However, in many cases the link between social and non-social symptoms is not clearly specified, for example, arguing that socio-communication problems result in RRBIs via a withdrawal from the social world. And while it seems plausible that non-social impairments in orienting attention could lead to difficulties in attention switching in social contexts (i.e., JA) there is, as yet, no empirical evidence for this relationship early in development.

Evidence from adult neuropsychological studies has suggested that social processing in the brain might be somewhat independent of other cognitive mechanisms, such as those which underlie language and memory (Anderson, Bechara, Damasio, Tranel & Damasio, 1999; Blair & Cipolotti, 2000; Fine, Lumsden & Blair, 2001). Further, genetic studies have demonstrated that while both social and non-social ASD impairments are highly heritable, they are broadly independent factors, with different underlying genetic etiology (Ronald, Happé & Plomin, 2005). However, this evidence comes from children and adults, not infancy, and therefore does not necessarily imply that these behaviours are independent over development.

Interactive Specialisation (IS; Johnson et al., 2005) is a theory of functional brain development which can be applied equally to both the social and non-social domains. IS suggests that processing and attention biases influence the infant's experience, which in turn affects the developing pattern of brain specialisation. It thus offers a framework that could explain the developmental emergence of cross domain difficulties in ASD. Atypical biases in attention early in infancy lead to different patterns of experience, thus reinforcing the atypical trajectory of development and leading to the behavioural symptoms that characterise the disorder. For example, difficulties in disengaging attention would lead to a different experience of the world, with longer looking at fewer objects. This, in turn, leads to altered potential learning experiences and shifts the infant's entire developmental trajectory.

To understand how social and non-social behaviours are related, it is important to look for risk markers in both domains early in development, and to establish how they map onto later outcomes. However, to date prospective longitudinal studies have tended to focus on individual markers. This approach is more in line with the idea of a single causal factor, posited by the majority of theories of ASD. Developmental models, on the other hand, offer different potential explanations for the development of ASD (Elsabbagh & Johnson, 2010).

Cumulative models argue that multiple risk factors work in an additive way to cause ASD. In cascading models, on the other hand, the role of the developmental process is emphasised, with variability in trajectories resulting from interactions between genetic and environmental factors leading to a non-linear mapping from risk factors to outcome.

Given that ASD is defined by behavioural impairments across the social and the non-social domains, in this thesis the interactions between risk markers will be examined. SEM, a multivariate modelling technique, lends itself to this type of analysis, looking at the relationships between multiple measures over time. SEM allows complex patterns of relationships to be tested simultaneously, rather than running many individual analyses. Thus, to address questions of how different measures interact across development, there is a need for this type of statistical approach.

1.9 Conclusion

There are various theories of ASD, which explain the disorder at different levels. Emerging findings from prospective studies are beginning to provide some support for different aspects of social and non-social theories. However, at present there is a lack of formal statistical modelling approach to examine how social and non-social behaviours early in life relate to later outcomes. In this thesis, I will first set out the advantages of the statistical modelling approach (Chapter 2). I will then look at the relationship between an early measure of social behaviour (gaze following) and subsequent clinical ASD diagnosis and language outcomes (Chapters 3 & 4). Chapter 4 also explores how high-risk toddlers use social feedback to learn words. In Chapter 5, I will examine how this social gaze following measure relates concurrently and longitudinally with a non-social measure (disengagement of visual attention). Finally, I will examine how these social and non-social behaviours together relate to ASD outcome, and the implications of this for understanding the development of

ASD. The final discussion (Chapter 6) will situate the current findings in the context of the results from prospective studies discussed in this introduction chapter, as well as thinking about the limitations of the thesis, and ideas for extending this work in the future. The way the different studies in this thesis fit together is illustrated in Figure 1.1.

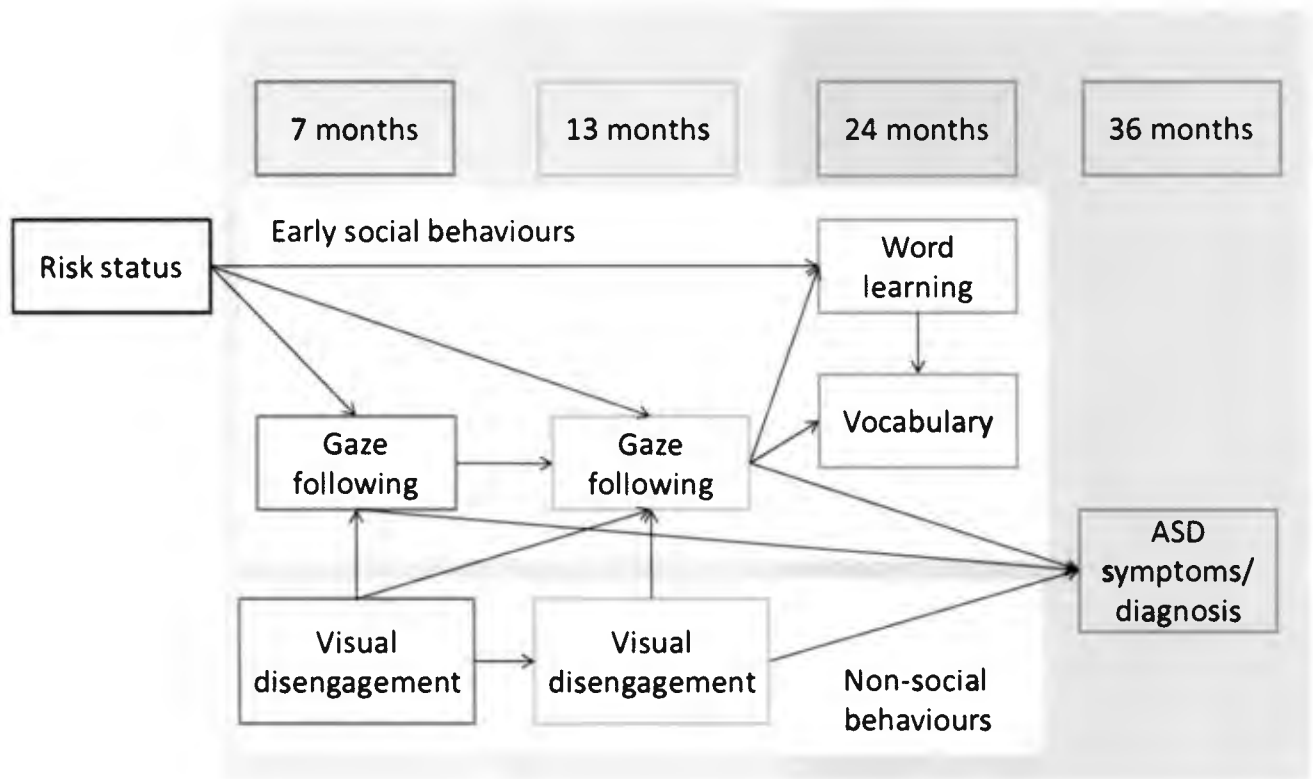


Figure 1.1 Relationships between behavioural measures over time, and their association with risk status and ASD outcome.

Overall, the aim of this thesis is to examine how social and non-social behaviours interact over development and their relationship with the emergence of autistic symptomatology, using structural equation modelling techniques to characterise developmental trajectories and model the relationships between multiple variables simultaneously. In order to address this overall aim, I will focus on the following key questions:

- 1) Does gaze following behaviour, a precursor to joint attention (JA), relate to risk group status and/or predict clinical outcome?

More specifically, two measures of gaze following behaviour were defined, infants' ability to follow the direction of another person's gaze shift (first look), and when gaze was correctly followed, the amount of time they then look to the object congruent with gaze direction (looking time). I aimed to test whether these two measures of gaze following were atypical in high-risk infants and particularly those who later developed ASD.

2) Does gaze following relate to concurrent and subsequent language development?

It was hypothesised that the looking time measure from the gaze following task would predict both a later parent report measure of receptive vocabulary and word learning ability following social feedback.

3) How do high-risk infants use social feedback to learn new words?

Here, I aimed to test for group differences (between low-risk and high-risk toddlers) on two measures: children's ability to make mappings between a heard word and an object using the mutual exclusivity principle (the principle that each object is referred to by a single name), and their ability to retain this mapping in long-term memory following feedback on their choice.

4) How does the combination of early social and non-social behaviours predict ASD?

First, I aimed to test whether a measure of social (gaze following) and non-social (visual disengagement) attention independently predict ASD outcome when included in the same model, and second, whether these measures relate to one another across the first year of development.

Summary of Chapter 1

- Early infancy measures can predict both typical and atypical outcomes.
- There are retrospective and prospective methodologies for investigating the early development of ASD.
- In the first year of life there are very limited behavioural symptoms of ASD, although there is some evidence for early differences in neural response to eye-gaze stimuli.
- From 12 months of age there are a number of social and non-social behavioural markers for the later development of ASD. However, these individual risk markers have typically been studied independently.
- There is a need for work looking at the effect of multiple markers on outcome. Structural equation modelling techniques lend themselves to the analysis of multiple variables.

Chapter 2

General Methods: An Introduction to Developmental Statistical Modelling

2.1 Introduction

This chapter will provide an overview of the methodology and statistical approaches taken in the studies presented in this thesis. The majority of developmental psychology research is currently dominated by the ‘static snapshot’ approach, due in part to the commonly available quantitative methods. At present, the data analysis methods typically employed constrain not only issues of design and analysis, but also conceptual development. The use of a more flexible modelling approach (e.g., Barnett et al., 2007; Gross, Shaw, Burwell & Nagin, 2009; Kerner & Muthén, 2009) offers a tool to address directly the challenges presented by longitudinal data.

The first part of the chapter will give a summary of the *British Autism Study of Infant Siblings* (BASIS) data set, and discuss the challenges associated with the analysis of longitudinal data. The second half of the chapter will introduce structural equation modelling (SEM) and use longitudinal data from the *Mullen Scales of Early Learning* (MSEL), a *standardised developmental assessment*, to demonstrate different models. A statistical model describes relationships between variables in terms of one or more equations. These models will address questions about developmental trajectories, inter- and intra-individual change, and the existence of subgroups of children within the high-risk group. The types of models presented in this chapter will be used in different studies throughout the thesis.

2.2 Participants

One hundred and four infants from the Phase 1 BASIS cohort took part in the studies presented in this thesis, 54 high-risk infants (21 male, 33 female) and 50 low-risk control infants (21 male, 29 female). Children were seen at the Centre for Brain and Cognitive Development when they were 6-10 months (*7 month visit*; $M = 7.35$ months, $SD = 1.21$) and 11-18 months (*13 month visit*; $M = 13.79$ months, $SD = 1.46$), 21-27 months (*24 month visit*; $M = 23.90$ months, $SD = 0.95$) and 32-53 months (*36 month visit*; $M = 37.93$ months, $SD = 3.02$). At the time of enrolment in the study, none of the infants had been diagnosed with any medical or developmental condition.

High-risk infants all had an older sibling (or in four cases, a half-sibling) with a community clinical diagnosis of ASD (45 male, 9 female). Diagnosis of the older sibling (henceforth, proband) was confirmed by two expert clinicians (PB, TC) based on information from the Development and Wellbeing Assessment (DAWBA; Goodman, Ford, Richards, Gatward & Meltzer, 2000) and the parent-report Social Communication Questionnaire (SCQ; Rutter, Bailey & Lord, 2003). The DAWBA is a parent-completed, web-based questionnaire that combines symptom ratings and narrative description that is then reviewed by an expert clinician. It was used to establish the prevalence of pervasive developmental disorders (ASD) in the UK national children and adolescent mental health survey (Fombonne, Simmons, Ford, Meltzer & Goodman, 2003). The SCQ is a parent-completed questionnaire with questions developed from the Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & LeCouteur, 1994). The majority of probands met criteria for ASD on both the DAWBA and SCQ ($n = 44$). While a small number scored below threshold on the SCQ ($n = 4$) no exclusions were made, due to meeting threshold on the DAWBA and expert opinion. For 2 probands, data were only available for either the DAWBA ($n = 1$) or the SCQ ($n = 1$). For four probands neither measure was available, and so inclusion was based on parent-

confirmed local clinical ASD diagnosis at intake in these cases. Parent-reported family medical histories were examined for significant medical conditions in the proband or extended families members, with no exclusions made on this basis.

Infants in the low-risk group were recruited from a volunteer database at the Centre for Brain and Cognitive Development, Birkbeck. Inclusion criterion was a lack of any ASD within first-degree family members (as confirmed through parent interview regarding family medical history). All low-risk infants had at least one older sibling and in five cases, only half-siblings (28 male, 22 female). Screening for possible ASD in these older siblings was undertaken using the SCQ, with no child scoring above instrument cut-off for ASD (≥ 15 ; one score missing).

Ethical approval was given by NHS NRES London REC (08/H0718/76) and parents gave informed consent.

2.3 Overview of the data set

At each of the four visits (7, 13, 24 and 36 months), children took part in a battery of assessments, including experimental tasks and standardised clinical and developmental assessments. Parental report questionnaire and interview measures were also collected (see Appendix 1 for the full BASIS network protocol). Here, only the measures which form part of this thesis will be discussed.

2.3.1 Clinical and developmental assessments

Mullen Scales of Early Learning (MSEL; Mullen, 1995): 7, 13, 24 and 36 month visits

The *MSEL* is a standardised developmental assessment, which examines early motor and cognitive development from 0-68 months. The assessment is comprised of five subscales: gross motor (GM), visual reception (VR), fine motor (FM), receptive language (RL) and expressive language (EL). An Early Learning Composite standard score (ELC) is calculated

for each child based on all scales apart from GM. In this thesis, the *MSEL* is used as a covariate to control for the effects of group differences in developmental level.

Autism Diagnostic Observation Schedule-Generic (ADOS-G, Lord et al., 2000): 24 month (high-risk only) and 36 month visits

The ADOS-G is a semi-structured observational assessment used in the diagnosis of ASD. The ADOS-G tasks involve social interaction with an examiner through play or conversation (depending on the age of the child) and a range of different behaviours, including language, gestures, eye contact, play and creativity and stereotyped behaviours are coded. The codes are 0, 1, 2 (and in some cases 3) with a higher score indicating a greater level of autistic-like atypicality. Certain item scores make up the final ‘algorithm’ scores, which are split into subsections: social, communication, creativity and repetitive and stereotyped behaviours. For the social and communication subsections of the algorithm, separate cut-off scores for ASD and autism are specified. Children scoring above the ASD/autism cut-off on both the social and communication subscales separately, and on the combined score meet criteria for the disorder.

Here, the ADOS-G was used (along with the ADI-R and expert clinical judgement) for the purpose of classifying the high-risk children into outcome groups, see below. Algorithm subscale scores at 24 and 36 months are also used as a measure of social-communication in Chapter 3.

Autism Diagnostic Interview-Revised (ADI-R): 36 month visit

The ADI-R is a structured parent interview, assessing triadic behaviours, which was used here solely for the classification of clinical subgroups (see below for more details).

Clinical outcome classification: 36 month visit

For the high-risk group consensus ICD-10 (World Health Organization, 1993) *ASD diagnoses* (ASD-sibs; childhood autism; atypical autism, other pervasive developmental

disorder, PDD) were achieved using all available information from all visits by experienced researchers (TC, KH, SC, GP). Given the young age of the children, and in line with the proposed changes to DSM-5, no attempt was made to assign specific sub-categories of PDD/ASD diagnosis. Toddlers from the high-risk group were considered *typically developing* (TD-sibs) at 36 months if they 1) did not meet ICD-10 criteria for an ASD; 2) did not score above the cut-off on the ADOS-G or ADI-R; 3) scored within 1.5 SD of the population mean on the *MSEL* Early Learning Composite (ELC) standard score (>77.5) and RL and EL subscale T-scores (>35). Finally, toddlers from the high-risk group were considered to have other developmental concerns, *atypically developing* (AT-sibs), if they did not fall into either of the above groups. That is, they either scored above the ADOS-G or ADI-R (Risi et al., 2006) cut-off or scored $<1.5SD$ on the *MSEL* ELC or RL and EL subscales.

From the 53 (out of 54) high-risk infants seen for diagnostic assessment at 36 months, 17 (11 male, 6 female) were classified as ASD-sibs (32.1%), 24 (7 male, 17 female) as TD-sibs (45.3%) and 12 (3 male, 9 female) fell into the AT-sibs group (22.6%), see Table 2.1. Of those classified as AT-sibs, 9 scored above ADOS-G cut-off for ASD, 1 scored above ADOS-G cut-off for ASD and below *MSEL* 1.5 SD cut-off, 1 above ADI-R cut-off, and 1 scored below *MSEL* 1.5 SD cut-off.

Table 2.1 Number of Infants in each Diagnostic Category

Low-risk		High-risk							
controls		infants		‘TD-sibs’		‘AT-sibs’		‘ASD-sibs’	
M	SD	M	SD	M	SD	M	SD	M	SD
n = 50		n = 54		n = 24		n = 12		n = 17	
(21m, 29f)		(21m, 33f)		(7m, 17f)		(3m, 9f)		(11m, 6f)	

2.3.2 Experimental measures

Using ‘infant looking’ as a window into cognition

Moll and Tomasello (2010) note that ‘until fairly recently, young infants were thought to be as cognitively incompetent as they were morally innocent’. Developmental researchers face the obvious methodological challenge that infants are unable to communicate through language, and it is therefore necessary to rely on interpretation of their behaviour. As discussed in Chapter 1, preferential looking and habituation paradigms have been used for over 50 years to test infants’ discrimination abilities. The consistent use of looking time measures to assess infant cognition is primarily due to the fact that infants can see immediately after birth and gain considerable control over their orienting responses within the first few months of life (e.g., Hood & Atkinson, 1993; Johnson, 1990). This is in contrast with gross motor behaviours, which infants gain control over much later, with actions such as coordinated reaching developing towards the end of the first year of life. The relatively early maturity of the visual system thus makes looking time an excellent candidate for investigating development in infancy.

Baillargeon and colleagues (Baillargeon, 1987; Baillargeon, Spelke and Wasserman, 1985) demonstrated that infants look longer, i.e., were more surprised, when shown an impossible event (e.g., a screen rotating through a solid box) compared to a possible one. These so called violation-of-expectation paradigms indicated that 3- to 5-month-old infants not only understood characteristics of the physical world, but that they could hold object representations in their memory. However, it has been argued by some researchers that such interpretations of infants' looking behaviours infer too much understanding, and in fact such looking patterns may alternately be accounted for by low level factors such as perceptual discrimination (e.g., Haith, 1998). Further, the exact same behaviour may be mediated by different underlying cognitive processes, for example, both blank staring at the screen and active looking would result in increased looking time, but are likely to reflect very different processes (Aslin, 2007).

More recently, research into infant visual attention has taken a slightly different approach, in which understanding underlying neural mechanisms, rather than cognition, is the primary aim (e.g., Colombo, 2002). Johnson et al. (1991) showed that infant looking behaviour, rather than being a unitary construct, could in fact be decomposed into orienting, engagement, maintenance, disengagement and shifting of attention.

In this thesis, two studies are presented which use infants' eye movements as the dependent variable. They fall neatly into the above framework, with looking time data from the *gaze following task* being used to make interpretations about infant understanding, and the *gap-overlap task* focused more on low-level visual attention measures. In Chapter 5, the relationship between these two measures and how they relate to ASD outcome is examined.

Gaze following task: 7 and 13 month visits

The aim of the gaze following task is to look at infants' ability to follow the eye-gaze of another person, in order to correctly fixate one of two objects. This task is based on a study

by Senju and Csibra's (2008), and used eye-tracking to record infants' gaze following behaviour longitudinally, at 7 and 13 months. Infants were shown short videos of a female model turning to look at one of two objects. Eye-tracking enables gaze following to be decomposed into different components. Measures of looking time to the congruent (gazed-at object) and first looks to the congruent versus incongruent objects were analysed.

The gap-overlap task: 7 and 13 month visits

The gap-overlap task is designed to look at infants' ability to flexibly shift and reorient their visual attention. In this task a central stimulus is initially presented, followed by the onset of a peripheral one. Saccadic reaction time is used to assess *disengagement* (speed of disengaging attention from the central fixation and orienting to the peripheral one when it appears).

Experimenter-child interaction studies

Piaget (1952, 1954) was perhaps the first to comprehensively document child development, based on observations and experiments with his own children. His theory stressed the concept of different stages of development. In the preoperational stage, which spans toddlerhood to children aged 7 years, Piaget argued that children develop linguistic competence but are not yet able to manipulate information in a logical way. While Piaget's theory represented a ground-breaking step forward in the way we think about cognitive development, the abilities of young children were still underestimated.

When considering the strategies employed by young children learning new words, there is still much debate in the literature over the interpretation of children's level of understanding. One such strategy is the mutual exclusivity (ME) principle, which states that children are biased towards preferring an object to have a single label (Markham, 1991). Carey and Bartlett (1978) asked 3-year-olds in a nursery classroom with two different coloured, but otherwise identical trays to get "the chromium tray, not the red one". After this single

labelling episode, a week later half the children had learnt something about the word chromium. That young children are able to use ME to disambiguate the referent of a new word is not disputed, however, researchers are divided over the interpretation of the cognitive process underlying this strategy. It may be that children use a one word per object, a lexical constraint. In other words, they already have a label for the familiar object (e.g., spoon) so the new word must refer to the novel object. However, the social-pragmatic account argues for a 'richer' interpretation, if the experimenter wanted the spoon he/she would have used the word 'spoon' because that is the word we both use for it, so he/she must be referring to the novel object.

Mutual exclusivity task: 24 month visit

The mutual exclusivity task is a live word learning task based on Horst and Samuelson's (2008) paradigm. Objects were presented on a tray and children were asked for a novel object in the presence of two familiar distractors and then, after a short delay, their memory for the novel object name was tested. This study looked at 1) children's ability to fast-map words using the mutual exclusivity principle; and 2) their use of ostensive cues in long-term retention of the word-object mapping.

Parental questionnaire methodology

MacArthur-Bates Communicative Development Inventory (CDI; Fenson et al., 1993): 7, 13, 24 and 36 month visits

The CDI is a standardised parent report measure of vocabulary and there are two versions used in this thesis, *words and gestures* (7 and 13 month visits) and *words and sentences* (24 and 36 month visits). In the first half of both versions of the questionnaire, parents fill in the number of words understood, or understood and said for words in different semantic categories, i.e., action words, people etc. The *words and gestures* version has a total of 396 words, and the second half of the questionnaire asks about communicative and symbolic

gestures that the child uses. The *words and sentences* questionnaire, on the other hand, has 655 words in the vocabulary section. The second part asks about the child's use of grammar and for three examples the child's longest sentences.

2.4 Longitudinal data analysis

Longitudinal data arises from repeated observations on one or more variables over time. This repeated measures nature of longitudinal data makes it more expensive to collect than cross-sectional data both in terms of the increased number of measurements (at least two for each participant) and the cost of tracking participants over time. However, for addressing questions in development, longitudinal data provides the only real way to examine the occurrence of change over time. Further, unlike cross-sectional data, which allows only within time correlations to be examined, longitudinal data enables the ordering of correlations over time to be assessed and thus comes closer to addressing questions of causality.

Despite the clear advantages of longitudinal data when addressing questions of development, there are several methodological issues which need to be taken into account. In order to overcome many of these difficulties certain assumptions need to be made and the validity of these assumptions will influence the accuracy and generalisability of the results. Bijleveld and van der Kamp (1998) give the example of measuring 'religiousness' to demonstrate the different possible explanations for an observed change. There are two alternative approaches to examine how age affects religiousness: cross-sectional and longitudinal. If religiousness is measured cross-sectionally in 20-, 40- and 60-year-olds, and an increase in religiousness is observed in the older participants, this could be due to an age effect (people grow more religious as they get older), but alternatively results could be explained by a cohort effect (people's views differ depending on the time they were born). If the question is approached longitudinally, and a group of 20-year-olds are assessed at three

time points this circumvents the issue of cohort effects. However, the results could be explained in terms of a period effect in which there is an overall change in the population over the course of the longitudinal study. These alternative explanations are linked linearly by the equation: cohort + age = period.

There are four basic longitudinal designs (Bijleveld & van der Kamp, 1998): simultaneous cross-sectional studies, trend studies, time series studies and intervention studies.

Simultaneous cross-sectional studies sample different age groups but on the same occasion (like the cross-sectional religiousness example), making the assumption that there is no effect of cohort. Further, no intra-individual change can be assessed as the individuals are different at each time point. Trend studies involve drawing a random sample of participants *of the same age* from the population, on different occasions (for example, looking at the change in mathematics ability in 12-year-olds over a ten year period of time by testing a different group of 12-year-olds every year). Like simultaneous cross-sectional studies, information on intra-individual change is not available in this type of design. In a time series design, the same participants are followed at successive time points, making it possible to assess both inter- and intra-individual change over time. The BASIS study has a prospective time series design but time series data more generally can be either retrospective or prospective. Intervention studies are a final type of longitudinal design which represent a variation of time series studies, in which an intervention or treatment affects only participants in the experimental group.

As discussed, longitudinal data can enable us to get closer to addressing questions of causality when thinking about developmental processes. However, when analysing longitudinal data, there are some methodological considerations that can influence the reliability of findings, which need to be considered and where possible accounted for. For

example, measuring the same variable over time violates the assumption of independent observations, as observations are in fact serially dependent.

One important issue when thinking about change over time relates to how ‘change’ is computed. One approach is to calculate a simple difference score between time 1 (T1) and time 2 (T2). So, for example, if we want to look at change in language scores in high-risk and low-risk infants, we could compute a difference score for both groups separately and compare them. This is called *unconditional analysis* and it compares the mean change in language scores. Whilst this may seem like a simple solution to addressing the question of change over time, there are various difficulties with this unconditional difference score approach. Firstly, scores at the T1 are likely to be related to subsequent scores. The phenomenon of regression towards the mean results in those with initially high scores having lower scores at later time points, and vice versa. However, the opposite can also occur with those who score higher on a language measure at T1 being likely to increase more rapidly (as, for example, they have a greater foundation for learning new language). Secondly, when calculating a difference score it is often assumed that the variance is equal across all time points. However, it is common in longitudinal studies for the variances to fan out and increase over time, and although standardised variables can be used, then we can only assess relative change. Another problem relates to the fact that variables such as language score are measured with a degree of error, and thus the observed variable does not represent the ‘true’ language score at either time point. This means that the calculated difference score will not represent the ‘true’ difference. This issue can be tackled by using latent variables (unobserved variables or ‘constructs’) to account for measurement error and then looking at the difference score between the latent variables.

Another approach to the question of analysing change over time is *conditional analysis*, in which T2 language scores are regressed on T1 scores. This answers the slightly different

question of whether, given the same initial language score, the groups have the same expected increase in language scores by T2. The advantage of conditional analysis, apart from avoiding some of the problems discussed above in relation to unconditional analysis, is that the idea of T1 having a causal effect on T2 is captured by using regression.

When analysing longitudinal data, missing data is a common problem and can arise for a variety of reasons ranging from missing questionnaire items to attrition from the study. There are three categories of missing data identified by Little and Rubin (1987, 1989): missing completely at random (MCAR), missing at random (MAR) and non-ignorable missing. MCAR is the easiest type of missing data to account for, but in reality the assumptions of MCAR do not often hold. When missing data are completely random the ‘missingness’ does not depend on the values of either observed or latent variables. When data is MCAR it is possible to just analyse full data cases. This is the strategy adopted by the majority of research in psychology using SPSS (Statistical Package for the Social Sciences) computer program with listwise deletion. However, in reality missing data is rarely related neither to observed or latent variables, and so the assumptions of MCAR are typically not met. MAR assumes that missingness is related to observed, but not latent variables or missing values. For example, if gender influences missingness of IQ score, but within males and females separately IQ score itself does not relate to the missingness of IQ data then for a model including gender the data are MAR. Although the assumptions of MAR are weaker than MCAR, they are more likely to hold in reality when missing data were not part of the study design (although it is not actually possible to test for MAR because the missing data scores are unknown). The final type of missing data, and the most serious in terms of implications for the results if listwise deletion is used, is non-ignorable missing data, which occurs when the missingness of data can be related to both observed but also to missing values and latent variables. An example of non-ignorable missingness would be if those with higher depression

are less likely to have missing data (because they were less likely to come in for the study). In this final case listwise deletion and standard full information maximum likelihood biases the results by systematically excluding participants.

In summary, longitudinal data are important in enabling questions of development to be directly addressed. However, such data are difficult and expensive to collect, as well as presenting methodological challenges from an analysis point of view. SEM offers a useful method for analysing longitudinal data as it allows the temporal dependence between measures to be explicitly modelled.

2.5 Structural equation modelling

The application of SEM techniques to sociological research questions (Blalock, 1964; Simon, 1954) first began over 50 years ago. SEM is a type of multivariate data analysis which can include both observed and latent (unobserved) variables. What a structural equation model is doing, in essence, is splitting the variance-covariance matrix, which is assumed to follow a multivariate normal distribution, among the relationships specified in the model. These relationships can be among the observed variables, among the latent variables and between observed and latent variables.

Structural equation models can be separated into two main constituents, the structural model and the measurement model. The structural model refers to the relationships between latent variables, as well as the relationship between independent observed variables and dependent latent variables. The measurement model on the other hand refers to the relationships between observed variables, and latent variables as independent variables predicting the observed variables. When the model only contains observed variables it is called a path analysis. Figure 2.1 shows the standard notation used in SEM to represent latent and observed variables and the relationships between them.





	Observed variable
	Latent factor
	Effect of one variable on another (e.g., regression)
	Covariance or correlation

Figure 2.1 Notation for structural equation models.

The modelling approach involves both estimating parameters in the model, and assessing the model ‘goodness of fit’. Goodness of fit refers to how well the proposed model recreates the variance-covariance structure observed in the data. Goodness-of-fit statistics can refer either to the absolute fit of the model or the relative fit of different models. Absolute fit statistics measure how similar the model generated data are to the observed data.

Comparative indices on the other hand assess the relative model fit between different models.

Perhaps, the most commonly referred to absolute test statistic is the χ^2 test of model fit, which represents the difference between the unrestricted (or observed data) covariance matrix and the restricted (or model) covariance matrix. If this test is not significant then we have no evidence to reject the null hypothesis, that there is no difference between the unrestricted and restricted models. The higher the significance value, the closer the fit between the data and the model. However, when interpreting this statistic it is important to consider sample size, as χ^2 is particularly sensitive to this. Given a large sample, the χ^2 value may often be significant even when the model does provide a good fit to the data (Jöreskog & Sörbom, 1993).

However, of more relevance to the models presented in this thesis is the fact that with small

sample sizes almost all models will fit because the test lacks power (see Satorra & Saris, 1985). Given these problems with the χ^2 test statistic, it is also necessary to consult other tests of model fit to look for converging evidence.

The other fit statistics that will be presented in this thesis when assessing model fit are the Comparative Fit Index (CFI; Bentler, 1990), the Bayesian Information Criterion (BIC; Raftery, 1993; Schwartz, 1978) and Root Mean Square Error of Approximation (RMSEA; Browne & Cudeck, 1993). The CFI has values ranging from 0 – 1 and for a good model fit values should be above at least 0.9 (Bentler, 1992) and more recently it has been argued that the value should > 0.95 (Hu & Bentler, 1999). The BIC takes into account not only model fit, but also the relative complexity of the model and is useful when comparing between models for the most parsimonious, and the lower the value, the better the fit (Byrne, 2012). Finally, the RMSEA is an absolute test of model fit and values < 0.05 indicate a good fit. When there is a small sample size, rejecting a true model is more likely to occur based on the RMSEA.

Once a model is found that fits the data, then we have one possible representation of the ‘causal structure’, or the relationship, among the variables (Bentler, 1980). In the majority of cases, model specification is a theory driven process, and deciding which relationships to include depends on prior hypotheses. In longitudinal data, the causal pathways are often dictated by the temporal sequence of measurement, e.g., we would expect an infant’s motor ability at 6 months to predict their motor ability at 12 months, not the other way around. In SEM, the relationships specified between variables can be represented using path diagrams or equations, although in this thesis I will only use path diagrams when presenting the different model types. Constrained or fixed pathway coefficients are represented in red, and estimated coefficients in black. Where possible STDYX standardised values are presented on the figures, and unstandardized values are used when these are not available. Grey arrows are

used to represent pathways with a non-significant coefficient and black arrows for pathways with a coefficient which is significant at $p < 0.05$ level.

The most basic type of SEM is a regression of a dependent variable on an independent variable (see Figure 2.2). In Figure 2.2, the regression coefficient of 12 month motor score on 6 month motor score is represented by a regression coefficient ' β '. The residual, ' e ', represents the disagreement between the data and the model i.e., $\text{data} - \text{model} = \text{error}$.

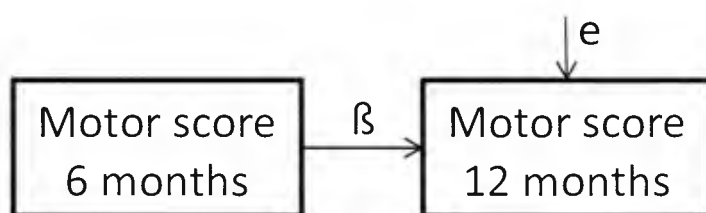


Figure 2.2 Regression of 12 month motor score on 6 month motor score.

This is the basic building block in an SEM model. Another key component in the majority of models run in this thesis is the inclusion of either 'group' which refers to risk status (i.e., high-risk versus low-risk controls), or 'outcome' which refers to 36 month outcomes (i.e., low-risk versus ASD-sibs, AT-sibs and TD-sibs). There are two possible strategies when modelling risk or outcome status, including group/outcome as time-invariant covariates, or a multiple groups approach, in which separate models are run for the different groups of children. There are advantages to running separate models for each group, namely checking measurement invariance across groups. However, given the limited sample size available in this study, group and outcome are included as time-invariant covariates in the model.

When using SEM, there are different modelling strategies that can be employed: strictly confirmatory, alternative models and model generating (Jöreskog, 1993). Strictly confirmatory analyses involve the specification of a single model based on the researcher's hypotheses. This model is then either rejected or not based on the data, but no further alterations to the model are made. The alternative models approach, which I use in the

models presented in this thesis, involves specifying several competing models, again based on hypotheses, and choosing the one that best represents the data (both from a statistical and theoretical point of view). In the model generating approach, once a pre-specified model has been rejected, exploratory changes to the model are made. For example, information about where the lack of model fit occurs is used to change the model specification.

Having decided on the modelling strategy and specified a particular model, estimation of the model can begin. Estimating structural equation models requires a relatively large sample size because an asymptotic estimator (e.g., maximum likelihood; ML) is used (MacCallum, Browne & Sugawara, 1996). For large sample sizes ML follows a χ^2 distribution under the null hypothesis i.e., that the covariance matrix is predicted by the model (thus a non-significant χ^2 test indicates a good model fit; Bollen, 1989). It is generally agreed that it is difficult to do SEM with a sample size much smaller than 100, and some argue that much higher sample sizes are ideally required (e.g., Boomsma, 1983 as cited in Bijleveld and van der Kamp, 1998). Having a small sample size can affect the reliability of the parameter estimates. This means that for data sets with only small sample sizes, simpler models, with fewer parameters, are a better option than more complex specifications.

In order to run a model it must be identified, i.e., all the parameters have only one solution. Although not necessarily sufficient in itself, it is a prerequisite for model identification that there are a fewer or equal number of parameters to be estimated than components in the variance-covariance matrix. This can be achieved either by fixing particular parameters to specific values, or by constraining them e.g., setting two parameters to be equal to each other. However, even when there are fewer free parameters than the information available from the variance-covariance matrix, it is still possible that the model may be unidentified empirically. This happens when the model estimate value for a particular parameter prevents the estimation of other parameters in the model (Rindskopf, 1984).

Having run a model which is identified and fits the data, the next step is to compare among several different models. For nested models, i.e., models obtained by constraining or fixing a parameter in a more complex model, the difference between the χ^2 goodness of fit test can be computed to see whether the simpler model provides a significantly worse fit than the more complex model. In the case that there is no significant difference between the fit of the two models, the more parsimonious model is typically chosen. In the final model, the significance values of the estimated parameters are examined. In this thesis, the standardised parameter estimates and associated significance levels will be reported. The standardised output assists in the comparison of the importance of different parameters.

2.6 Data analysis: the Mullen Scales of Early Learning

In this section, using data from the *MSEL* (see Table 2.2) I will present several different modelling approaches, including path analysis and latent variable modelling and discuss their relative advantages and disadvantages. Further, I will make comparisons between these models and the analyses such as analysis of variance (ANOVA), regression and t-tests, which are more typically used at present to address questions in the field of developmental psychology (e.g., British Psychological Society conference, Developmental Section, 2011).

Table 2.2 Descriptive Statistics for the Mullen Scales of Early Learning (MSEL) Subscale

Standard T-scores and Early Learning Composite Standard Scores (ELC)

	Low-risk controls		High-risk infants		'TD-sibs' M		'AT-sibs' M		'ASD-sibs'	
	M	SD	M	SD	SD		SD		M	SD
7m MSEL	n = 50		n = 53		n = 24		n = 12		n = 16	
ELC	104.42	11.31	94.00	12.88	96.13	11.77	92.83	8.13	92.13	17.30
	n = 50		n = 53		n = 24		n = 12		n = 16	
GM T-score	50.40	9.09	45.40	9.99	45.17	8.39	45.67	10.08	46.06	12.58
	n = 50		n = 53		n = 24		n = 12		n = 16	
VR T-score	54.32	8.70	50.53	8.96	51.71	8.28	48.58	6.90	50.00	11.49
	n = 50		n = 53		n = 24		n = 12		n = 16	
FM T-score	58.00	9.41	52.45	10.47	53.54	11.11	53.92	8.76	49.81	11.08
	n = 50		n = 53		n = 24		n = 12		n = 16	
RL T-score	46.04	9.04	41.42	11.80	44.79	10.38	39.08	8.37	38.81	15.08
	n = 50		n = 53		n = 24		n = 12		n = 16	
EL T-score	50.20	8.54	42.98	7.07	41.67	4.78	43.75	7.47	44.69	9.45
13m MSEL	n = 47		n = 53		n = 23		n = 12		n = 17	
ELC	106.11	15.73	97.40	17.91	103.35	18.12	99.75	11.35	89.18	18.30
	n = 48		n = 53		n = 23		n = 12		n = 17	
GM T-score	51.06	16.08	46.26	16.55	48.83	15.53	43.92	16.36	45.53	18.44
	n = 47		n = 53		n = 23		n = 12		n = 17	
VR T-score	55.70	9.39	51.19	10.55	52.35	11.98	55.17	5.51	47.71	10.11
	n = 47		n = 53		n = 23		n = 12		n = 17	
FM T-score	61.17	9.10	54.94	12.44	56.74	12.45	58.25	10.64	51.18	12.91
	n = 48		n = 53		n = 23		n = 12		n = 17	
RL T-score	46.10	12.36	43.36	12.95	48.04	13.83	43.92	10.73	37.76	10.69
	n = 48		n = 53		n = 23		n = 12		n = 17	
EL T-score	48.08	9.36	44.42	11.46	49.17	11.24	41.83	9.10	39.88	11.70
24m MSEL	n = 42		n = 52		n = 24		n = 12		n = 16	
ELC	116.02	13.98	102.25	19.77	105.38	17.53	102.00	16.84	97.75	24.74
	n = 37		n = 37		n = 17		n = 8		n = 12	
GM T-score	59.89	8.08	45.19	11.24	44.59	11.35	48.25	8.17	44.00	13.21
	n = 45		n = 52		n = 24		n = 12		n = 16	
VR T-score	58.93	12.42	53.17	12.92	56.21	10.48	49.17	14.91	51.63	14.35
	n = 43		n = 52		n = 24		n = 12		n = 16	
FM T-score	54.33	8.75	49.94	9.37	51.21	8.46	51.17	8.85	47.13	10.91
	n = 43		n = 52		n = 24		n = 12		n = 16	
RL T-score	59.14	7.84	50.50	14.46	53.42	13.99	51.83	8.07	45.13	17.81
	n = 44		n = 52		n = 24		n = 12		n = 16	
EL T-score	58.16	11.70	49.67	13.30	49.17	12.58	49.92	11.47	50.25	16.17
36m MSEL	n = 48		n = 52		n = 24		n = 12		n = 16	
ELC	115.77	16.25	105.38	21.52	113.46	13.26	103.42	18.98	94.75	28.51
	n = 48		n = 37		n = 24		n = 12		n = 16	
VR T-score	59.00	10.91	56.04	13.87	60.17	9.13	55.67	12.41	50.12	18.66
	n = 48		n = 52		n = 24		n = 12		n = 17	
FM T-score	56.58	13.93	48.58	15.49	54.00	13.70	50.67	14.70	39.47	15.05
	n = 48		n = 52		n = 24		n = 12		n = 17	
RL T-score	57.40	10.22	50.94	12.19	54.13	6.91	50.42	7.79	46.82	18.42
	n = 48		n = 52		n = 24		n = 12		n = 17	
EL T-score	58.98	9.12	52.91	13.14	58.42	9.49	49.25	12.80	47.71	15.32

The distributions of the *MSEL* Early Learning Composite scores are presented in Appendix 2. While the latter two visits appear slightly negatively skewed, the tests of skewness and kurtosis were within the normally distributed bounds, so no transformations were applied. The T-scores for all the individual subscales were also normally distributed.

Models presented throughout the thesis are run using Mplus (see Appendix 3 for model scripts), STATA or SPSS computer programs. STATA and SPSS are both software packages that can be used for a range of basic statistical analyses. Mplus is a latent variable modelling program which allows flexible models including both categorical and continuous observed and latent variables. Figure 2.3 (from the Mplus website:

<http://www.statmodel.com/features.shtml>), shows the different possible models that can be run using this program.

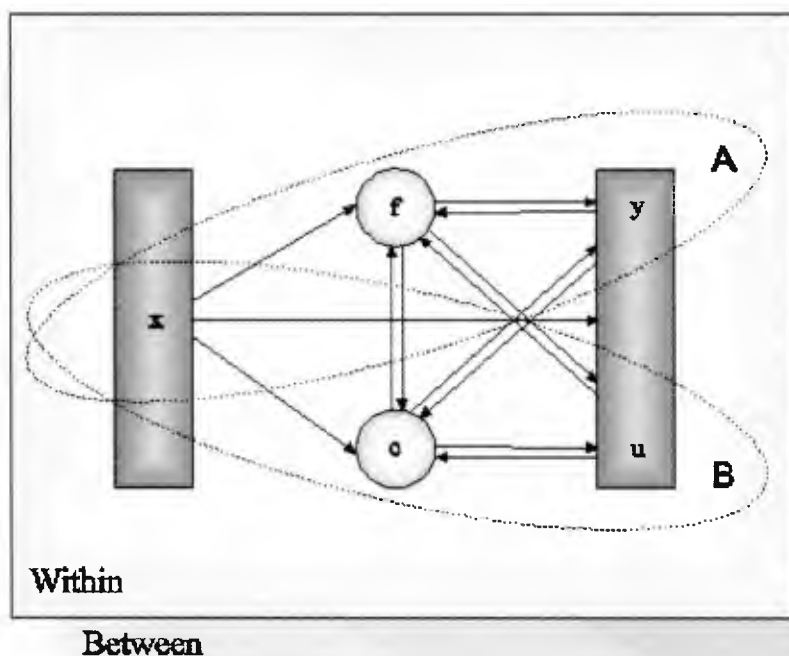


Figure 2.3 The Mplus framework.

In Figure 2.3, the circles represent latent variables and the rectangles observed variables. ‘f’ stands for factor, i.e., a continuous latent variable and ‘c’ stands for class, a categorical latent variable. Similarly ‘y’ and ‘u’ represent continuous and categorical observed variables

respectively. The arrows specify the different possible relationships that can be modelled between these variables, and the within and between rectangles reflect the fact that multilevel models can be run in this framework.

2.6.1 Autoregressive models

Autoregressive models work by analysing how the variance-covariance matrix of variables changes over time (e.g., see Figure 2.4a). This type of model, with only direct effects from one time point to the next is called a *first order autoregressive model*. Time 1 (i.e., 7 month visit) can directly affect the score at time 2 (i.e., the 13 month visit) but only indirectly affect the score at time 3 (T3; i.e., 24 month visit), via the effect at 13 months. *Higher order autoregressive models* on the other hand, contain a direct pathway from T1 to T3 (7 month – 24 month visits), thus enabling score at 7 month to influence the 24 month score both directly, and indirectly via 13 month score (see Figure 2.4b).

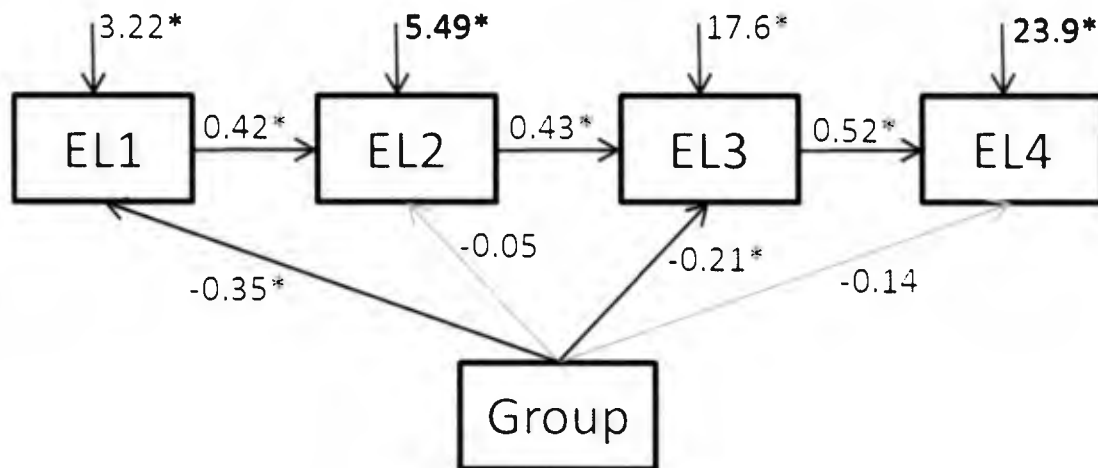


Figure 2.4a First order autoregressive model for raw expressive language (EL) score at 7 – 36 month visits.

The model in Figure 2.4a provides a good fit to the EL data ($\chi^2(3) = 1.72$, $p = 0.63$, CFI = 1.0, RMSEA < 0.001). Results from the STDYX standardisation (which standardises on

both X and Y variables) show that, after controlling for the effect of group, all the ‘autoregressions’ are significant, i.e., 7 month EL predicts 13 month EL ($\beta = 0.42$, standard error; S.E. = 0.09, $p < 0.001$), 13 month EL predicts EL at 24 months ($\beta = 0.43$, S.E. = 0.08, $p < 0.001$) and 24 month EL predicts EL at 36 month ($\beta = 0.52$, S.E. = 0.08, $p < 0.001$). According to this model, 33.6% of the variance in 36 month EL is accounted for by the direct effect of 24 month EL and group, and the indirect effects of EL at 7 and 13 months.

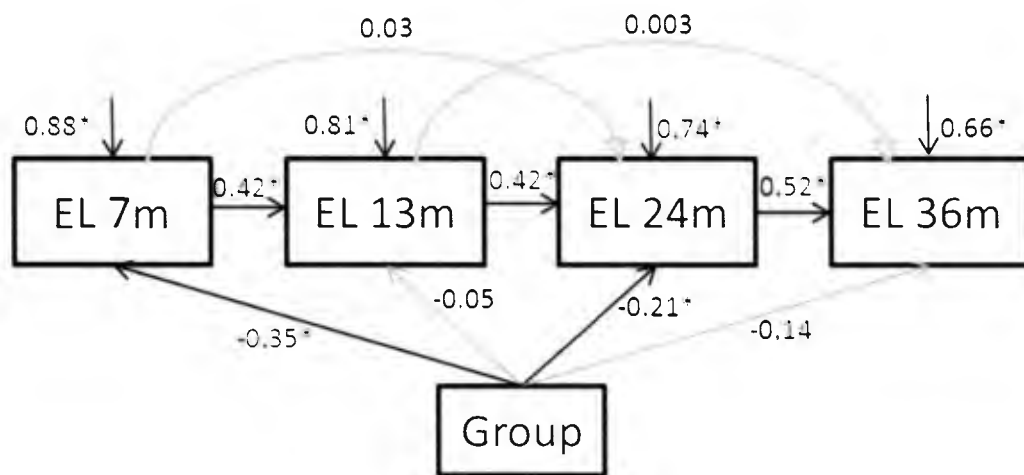


Figure 2.4b Second order autoregressive model for raw EL score at 7 – 36 month visits.

The model in Figure 2.4b is a second order model, with direct pathways now included from 7 to 24 months, and 13 to 36 months of age. The model provides a reasonable fit to the data ($\chi^2(1) = 1.66$, $p = 0.20$, CFI = 0.99, RMSEA = 0.08) although the RMSEA is a little higher than the ideal value of < 0.05 . Similarly to the previous model, 7 month EL predicts 13 month EL ($\beta = 0.42$, S.E. = 0.09, $p < 0.001$), EL at 13 month predicts 24 month EL ($\beta = 0.42$, S.E. = 0.09, $p < 0.001$) and 24 month EL predicts EL at 36 month ($\beta = 0.52$, S.E. = 0.09, $p < 0.001$). However, there are no significant pathways between EL at non-adjacent time points, so the first model gives a more parsimonious account of the data. This suggests that, having controlled for the effect of group, EL score at each time point strongly predicts

EL score at the subsequent time point, with only indirect relationships between EL scores at non-adjacent time points.

Another type of autoregressive model uses latent variables to account for measurement error. The latent variable here represents 'true' language score (see Figure 2.4c). This is a more complex model and in order for it to be identified, certain assumptions must be made. Here, factor loadings have been fixed to 1, the factor covariances fixed to 0, and the residual variances of EL constrained to be equal. Further, in this case, unlike models 2.4a and 2.4b, in the best fitting model only the first factor (i.e., true language score at time 1) was been regressed on group. This may be the case here, unlike in the previous models, because of the nature of latent autoregressive models. In a basic first-order autoregressive model, variability consists of both true score variance and occasion specific error, the latter of which is not transmitted in the coefficients. This means that the coefficients are smaller than they should be and so rather than the effect of group being transmitted via the autoregressions, direct pathways from group to language at later time points become significant. However, in this latent model, the coefficients are larger as there is only true score variance. This means that the effect of group here is transmitted through the regression coefficients, and does not also need to be accounted for directly at later time points.

This model provides a reasonable fit to the data ($\chi^2(5) = 8.94$, $p = 0.11$, CFI = 0.96, RMSEA = 0.09) although ideally the RMSEA should be lower. The factors are significantly indicated by EL at each time point. As shown in Figure 2.4c, all the autoregressions between the factors are significant (f2 on f1, $\beta = 0.64$, S.E. = 0.17, $p < 0.001$; f3 on f2, $\beta = 0.56$, S.E. = 0.11, $p < 0.001$; f4 on f3, $\beta = 0.60$, S.E. = 0.07, $p < 0.001$).

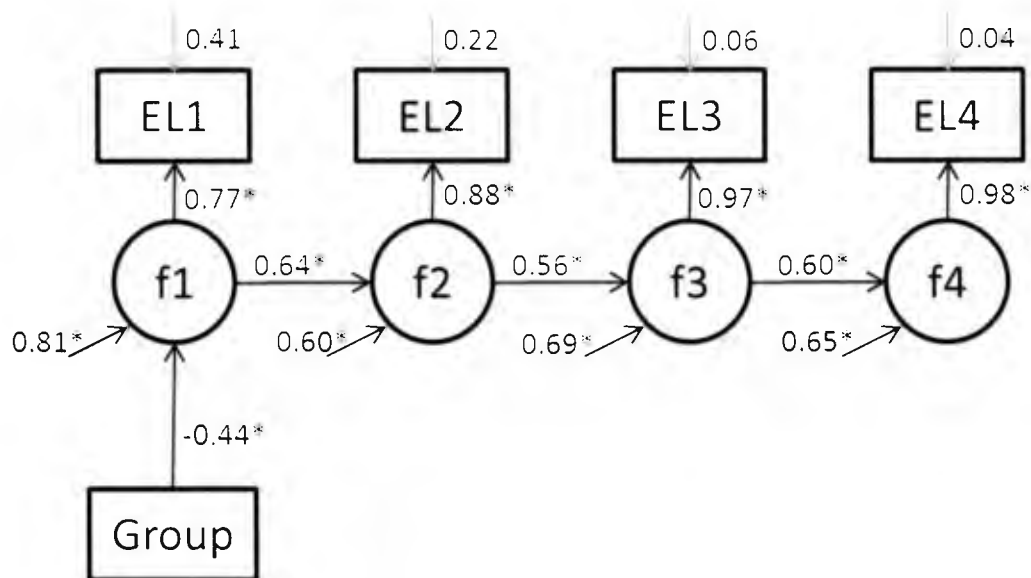


Figure 2.4c Latent variable first order autoregressive model for raw EL score at 7 – 36 month visits.

Given that the first-order model (Figure 2.4a) is the most simple and provided an excellent fit to the data on all fit criteria, this is probably the model that we should go with in this case, as it is the most parsimonious account of the data. However, all models show broadly the same pattern, with only direct influences between EL at subsequent time points and an effect of group on EL at 7 months, but a weaker effect at 13 month – 36 months.

2.6.2 Growth curve models

Another type of model useful for longitudinal data analysis is the growth curve model (GCM). GCMs are an important tool for examining inter- and intra-individual change over time. When data from at least three time points is available, it is possible to test whether linear growth fits the data, and with more time points, the fit of curvilinear functions to the data can be assessed. There are two different approaches to running GCMs, the latent modelling framework which has come from the SEM background, and multilevel GCMs which have arisen from a hierarchical linear modelling perspective. This latter approach is typically used for the analysis of nested variables, for example, children within classes within

schools, although it is also possible to analyse repeated measures data in this way (Bryk & Raudenbush, 1987) with time as a predictor. However, in this thesis the latent GCM approach will be taken, as it parameterises time, enabling a flexible approach in which time can be treated either as a categorical variable (i.e., 7, 13, 24 and 36 month visits) or a continuous variable (i.e., ages 7-36 months) simply by changing the regression coefficients that relate the observed variables to the slope.

In a GCM, children can differ from one another both on their score at T1 (intercept) but also on their subsequent rate of growth, from time 1 to time 4 (T4; slope). Here, the ‘intercept’ refers to the start of the GCM, i.e., score at T1, rather than the value of Y when X is zero, which is why the intercept in a GCM is also often referred to as the initial status (Muthén & Khoo, 1998). In a latent GCM both the intercept and slope of the regression equation are specified as latent variables, which are random and can vary across individuals. This is very different to linear regression where the model slope and intercept take a fixed value. The ability to quantify individual differences in developmental trajectories gives GCMs a substantial advantage over the techniques typically used by developmental psychology researchers (e.g., Curran & Hussong, 2003).

GCMs, like autoregressive models, explain change in the variance-covariance matrix over time (the intercept and slope each have a variance and associated covariance). However the description of the data using a random slope and intercept in a GCM is a conceptually more elegant or more parsimonious way to look at change over time. In a GCM, the mean of the intercept is fixed at zero. The regressions of observed variables on the intercept latent variable are constrained to be equal to one another, to reflect that fact that the same variable is measured at each time point (see Figure 2.5). So, using the expressive language example, constraining these pathways to be equal to one another reflects the EL score given no growth

between T1 and T4. While in theory any values could be used here, the convention is to use '1'.

The mean of the slope is freely estimated, and growth in the observed variables is represented by multiplying the mean of the slope by the regression coefficients from the slope to the observed variables. These coefficients are part of the variance-covariance structure, and are fixed to particular values. For example, 0, 1, 2, 3 would represent linear growth between equidistant time points. The pathway from T1 to the slope is fixed at '0' because there is no growth at T1 (as this is the initial status). Values 1, 2 and 3 reflect visits separated by equal time, for example, visits one year apart. For the *MSEL* data at 6, 12, 24 and 36 months, linear growth can be represented by 0, 1, 3, 5 (see Figure 2.5). The '0' pathway from slope to 7 month EL is left out, by convention.

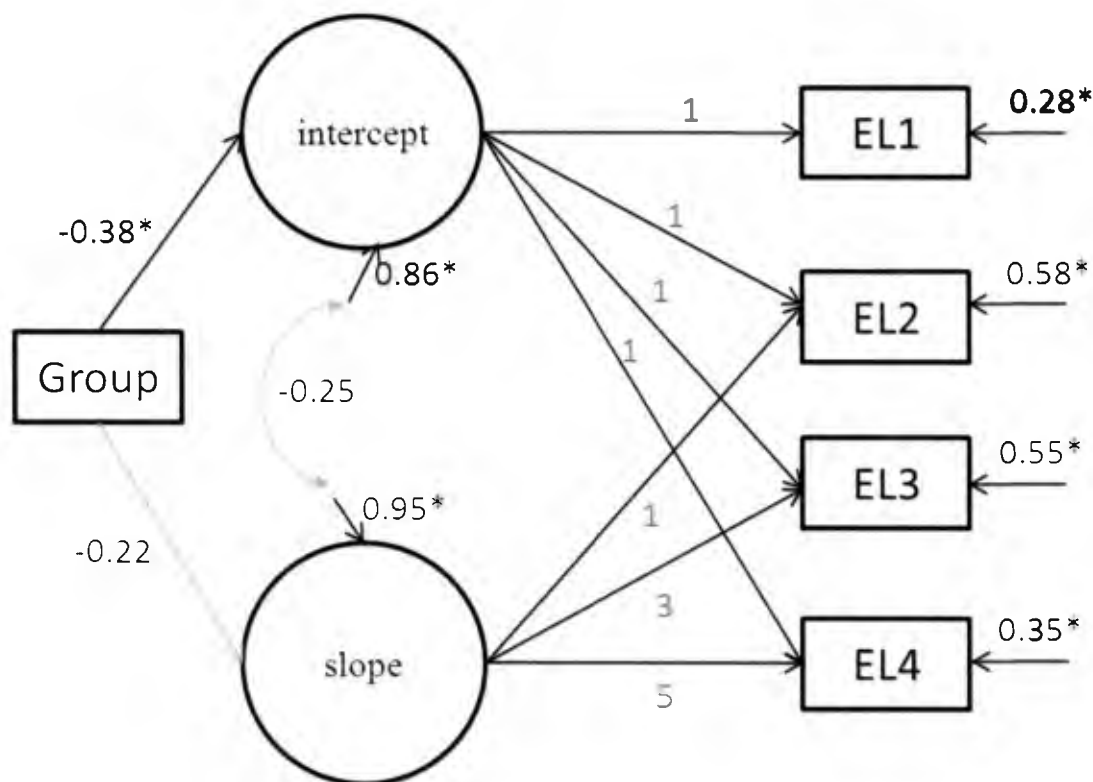


Figure 2.5 Example of a GCM using EL score at 7 – 36 months.

When this model was run, it provides a reasonable fit to the data (see Figure 2.6; $\chi^2(7) = 12.80$, $p = 0.08$, CFI = 0.94, RMSEA = 0.09) although the CFI is a little lower and the RMSEA a little higher than we would ideally like them to be. The output shows group to have a significant effect on the intercept (in this case EL score at 7 month visit; $\beta = -0.38$, S.E. = 0.10, $p < 0.001$), with high-risk infants having lower initial EL scores compared to low risk children. The effect of group on the slope (rate of EL growth from 7–36 month visit; $\beta = -0.22$, S.E. = 0.12, $p = 0.06$) did not reach significance, but the trend is in the expected direction with expressive language in high-risk children increasing more slowly than in low-risk children.

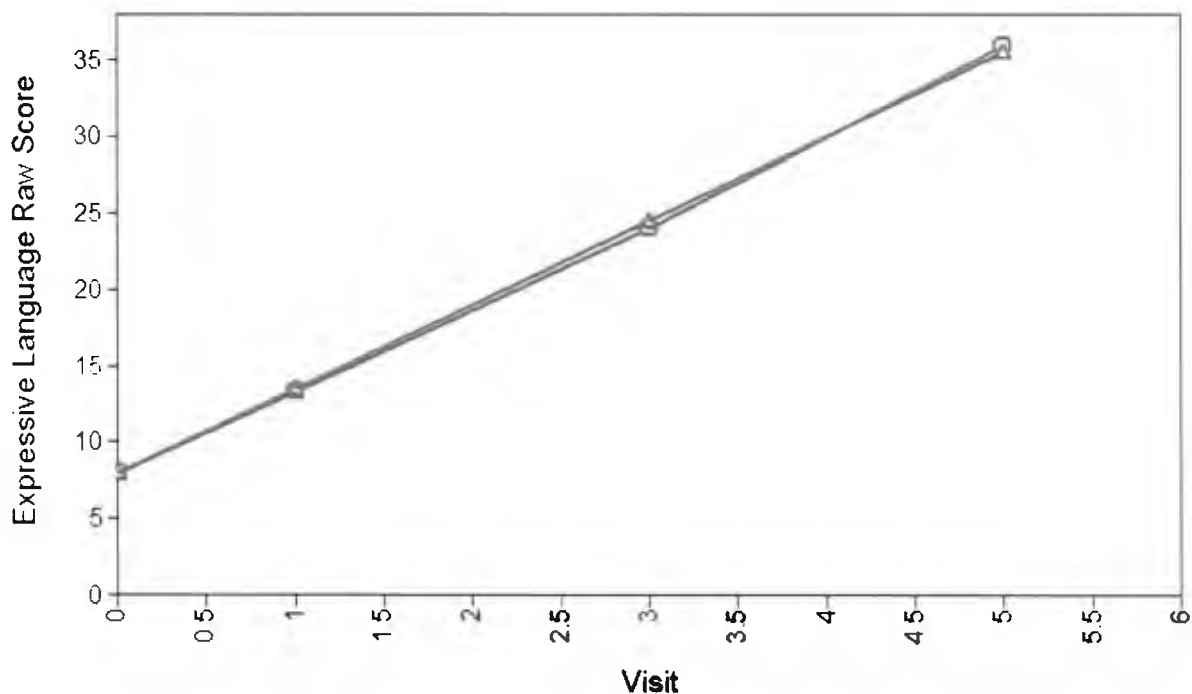


Figure 2.6 Graph of the sample (●) and model estimated (▲) means. Because the slope regression coefficients were set at 0, 1, 3, 5, these values on the x-axis correspond to the 7, 13, 24 and 36 month visits respectively.

When modifying this model, there are a variety of possibilities. It is possible to constrain the residual variances of EL to be equal at each time point. However, it is important to think about whether this is a valid assumption to make, given the particular model. For example, this may be less appropriate when modelling raw scores from a longitudinal study, as variability often increases over development and so measurement error is also likely to increase. Another possibility with four time points is to add a quadratic factor. In this case, the model fit statistics become: $\chi^2(2) = 3.48$, $p = 0.18$, CFI = 0.98, RMSEA = 0.08, which look better than the previous model with just an intercept and slope. However, we have lost 5 degrees of freedom, which means to accept this new model over the previous simpler one we need a drop in χ^2 value of 11.07. However, χ^2 only drops by 9.32 (12.80 - 3.48), so in fact this quadratic model is not a significantly better fit.

It is also possible to run a slightly different type of model that takes advantage of the fact that there was variability in the age at which children were tested. In fact, whilst the mean age at visit 1 was 7.35, infants ranged from 6-10 months. To directly utilise this variation in the model, it is possible to set the regression pathways to the age of the child (rather than 0, 1, 3, 5). Importantly, this changes the interpretation of mean of the intercept, because now it represents the average score when the child was '0' years old (somewhat less useful than the average score at visit 1).

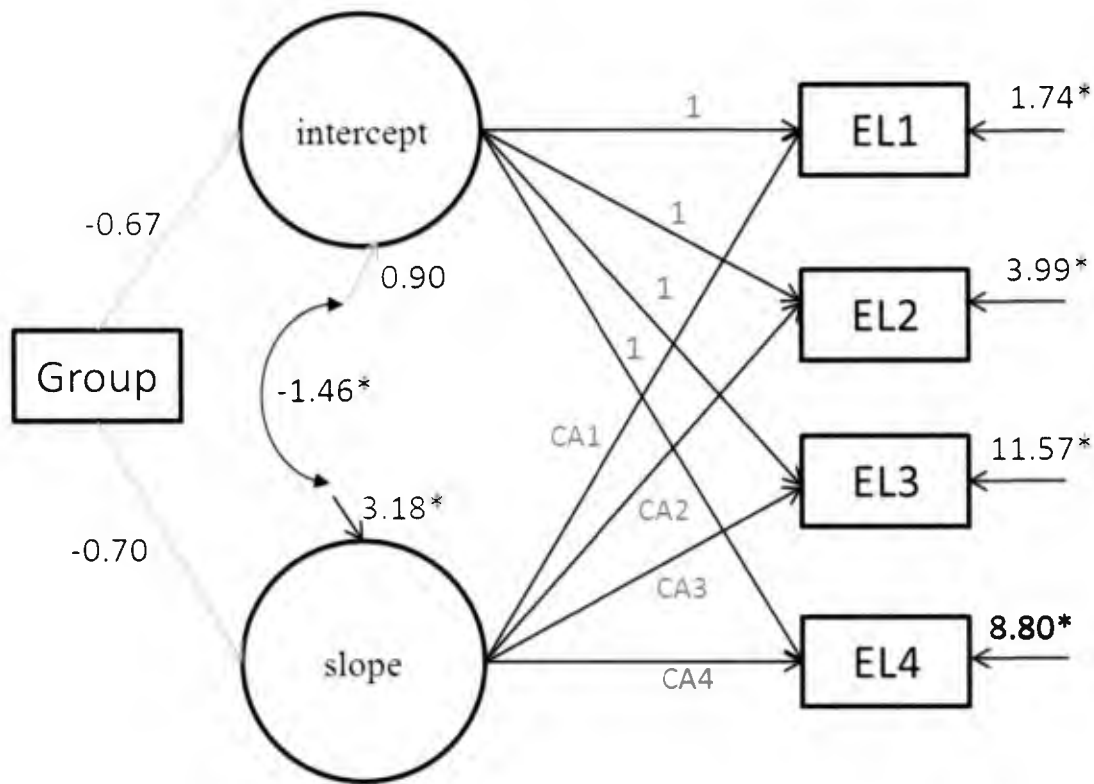


Figure 2.7 Example of a GCM using chronological age for EL score at 7 – 36 months.

When running this type of model in Mplus, the absolute fit statistics are not available. The BIC value is given, which allows for comparison between models, i.e., to assess whether adding a quadratic factor improves model fit. For the model in Figure 2.7, (BIC = 2004.8), the results show no significant effect of group on the intercept (which is now score at 0 years) or slope, although as before the direction of effects is in the expected direction. The coefficients here are unstandardised because STDYX is not available in Mplus when running this type of ‘random’ analysis. Interestingly, in this model the intercept and slope are significantly negatively correlated (covariance = -1.46, $p = 0.002$), which means that, having accounted for the effect of group, individuals who start with higher EL scores grow more slowly.

2.6.3 Growth mixture models

Unlike latent GCMs, which look at individual variation around a single mean curve, a growth mixture model (GMM) allows for heterogeneity in trajectories, and assumes that subgroups exist within a single population. This is particularly useful here, when we have a theoretically motivated reason for expecting different subgroups, based on outcome classification.

When running a GMM it is necessary to specify the number of latent classes, a process which should be informed both by theory and statistics. Here, models are started with two classes (which may represent the low-risk versus high-risk difference) and then the number of classes is increased, one at a time, up to four (which could reflect outcome groups, low-risk controls, TD-sibs, AT-sibs and ASD-sibs). The BIC value, which is a function of the likelihood that favours parsimony in the model, will be used to distinguish between different models. Further, the entropy value, which is related to certainty of class membership, should be above 0.8 for a good model.

To make sure the model is converging on the right solution random starts are needed. To understand why this is important imagine a class of outward bound students each skydiving out of a separate plane at different locations over the Alps (see Figure 2.8). Their objective, on landing, is to find the mountain nearest to them and climb to the top, in order to locate the highest point in the region. The highest point is assumed to be the one with the student (or group of students) who report the highest value. However, if there aren't enough people, then it is possible that the highest point reported by any student will actually be 'local maxima' i.e., a medium sized mountain. Increasing the number of random starts increases the likelihood that the true highest point will be found, and the number required depends on sample size and complexity of the model (i.e., number of classes and number of observed variables).

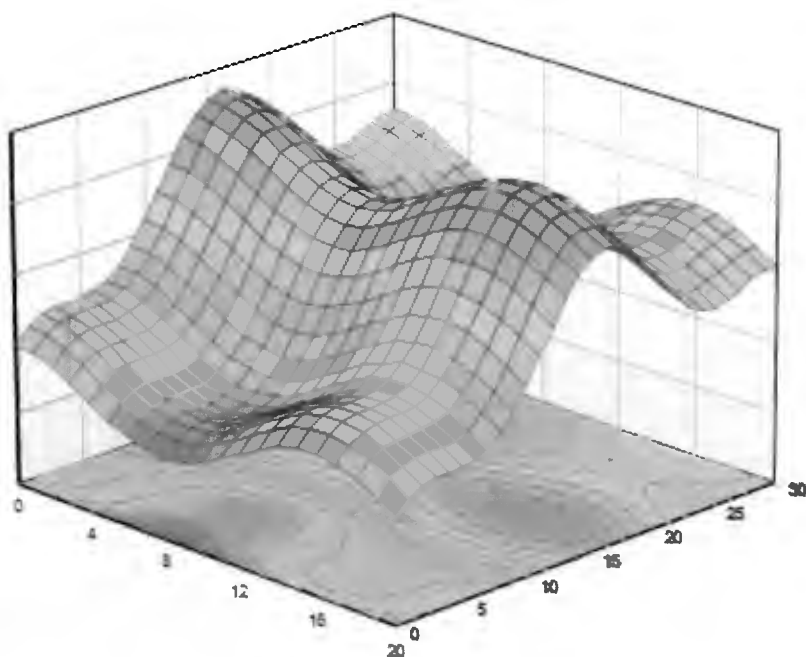


Figure 2.8 Local versus global maxima.

The limited sample size resulted in unstable models with a two- three- and four- class models each containing at least one very small class with < 10 individuals (see Appendix 4).

2.7 Summary of the methodological approach

Unlike more basic methods, using a structural equation modelling approach enables more complex questions to be addressed, such as how does the rate of expressive language development predict the rate of receptive language development, or are there subgroups of infants who show different developmental trajectories. These methods also allow some of the challenges associated with longitudinal data, such as missing data and floor and ceiling effects to be directly accounted for in the model. It is thus an appealing tool to address questions of development, particularly with infancy data, which is often subject to missingness, both at the overall visit level, but also in terms of number of trials completed.

Summary of Chapter 2

- Chapter 2 provided an overview of the methodological and statistical approaches taken in the studies presented in this thesis.
- There are important advantages as well as challenges associated with analysis of longitudinal data.
- Structural equation modelling offers a technique to deal explicitly with many of the challenges associated with longitudinal data, such as missing data.
- Examples of autoregressive, latent autoregressive, growth curve and growth mixture models were presented using data from the *MSEL*.

Chapter 3

The Gaze Following Task: Eye-tracking of Gaze Following Behaviour in Infants at High Risk for Autism Spectrum Disorder

3.1 Introduction

As discussed briefly in the Chapter 1, joint attention (JA) type behaviours are among the earliest discriminators of children who subsequently go on to develop an ASD. However, these behaviours do not emerge until towards the end of the first year of life, and little is known about the development of the precursor ability of gaze following. In this study, eye-tracking was employed to measure gaze following longitudinally, at 7 and 13 months, in high-risk infants and low-risk controls.

This chapter is split into two parts, with Part one focusing on defining and separating out early measures of gaze following and investigating whether high-risk infants, particularly those who later go on to develop ASD, show difficulties in their gaze following behaviour. The second part of this chapter examines the relationship between gaze following measures and language outcomes. It is well-known that links between JA and language exist in both typical development and in young children with ASD, but much less is known about how the earliest measures of gaze following behaviour in the first year of life might predict concurrent and subsequent vocabulary and language development. Here, I examine the relationship between gaze following at 7 and 13 months and parent reported vocabulary at 13 and 24 months (the *MacArthur-Bates Communicative Development Inventory, CDI*; Fenson et al., 1993).

3.2 The emergence of gaze following and joint attention in typical development

In typical development, sensitivity to another person's gaze appears to be present from immediately after birth. New-born infants show a preference for faces with eyes open (Batki, Baron-Cohen, Wheelwright, Connellan & Ahluwalia, 2000) and fixate for a greater amount of time, with a higher number of orienting responses, to faces with direct as compared to averted gaze (Farroni, Csibra, Simion, & Johnson 2002). Neonates can also be 'cued' by the direction of an adult's gaze, with faster orienting to a target congruent, rather than incongruent, with gaze direction (Farroni, Massaccesi, Pividori, & Johnson, 2004). These authors used an attention cueing paradigm in which they presented a central face, whose eyes blinked to gain the infant's attention, followed by a gaze shift to the right or left. A target stimulus then appeared on the screen in a location congruent or incongruent with the direction of gaze. Infants were quicker to orient (based on saccadic reaction time) to the target when it was congruent with gaze direction. Frischen, Bayliss and Tipper (2007) argued that infants are relying on fairly 'low-level' factors, such as the direction of movement of the pupil. However, the cueing effect disappears following inversion of the face or removal of preceding direct gaze (Farroni, Mansfield, Lai & Johnson, 2003), suggesting that it is something about object-directed motion within the context of an upright face with preceding mutual gaze that is important for gaze following to occur (see also Senju & Csibra, 2008).

Gaze following involves the orienting of attention towards a stimulus in response to another's shift in gaze. Not only is such gaze following behaviour present early in infancy, but it is also observed in other social primates (Tomasello, Call & Hare, 1998; Deaner & Platt, 2003). Further, there is evidence that even dogs are able to use eye-gaze and head

direction cues, to locate food, when these cues are not in conflict (Hare, Call & Tomasello, 1998). The fact that this ability is preserved across species suggests that gaze following may be a fairly low-level process. Butterworth and Jarrett (1991) argue that in typical human development, early emerging biases, such as sensitivity to eye-gaze and gaze following may form a mechanism for the subsequent development of joint attention (JA). JA has been defined as the ability to engage in *shared attention* with another individual (Baron-Cohen, 1989b; Mundy, Sigman, Ungerer & Sherman, 1986). The common examples of JA behaviours include spontaneous gaze following and ‘protodeclarative’ pointing (Scaife & Bruner, 1975). These qualify as JA because the child’s response brings them into a shared focus of attention with another person. It is not restricted to these since other gestures (e.g., a nod in one direction) can produce the same end-state (the other person turning to look at the same ‘topic’ picked out by the first person’s nod).

In this study a distinction is drawn between gaze following and JA. By my definition JA also implies referential understanding, whereas gaze following, when taken in isolation, does not. In their developmental model, Tomasello et al. (2005) argue that children move from the ‘understanding of pursuit of goals’ stage at 9 months to ‘understanding choice of plans’ between 12-15 months. They suggest that the key change is a switch from ‘joint perception’ (i.e., gaze following to a shared target) to ‘joint attention’, mediated by the development of a capacity to represent others’ internal mental representations. Whilst not all researchers take this modular approach, (e.g., Mundy, Sullivan & Mastergeorge, 2009) it certainly seems that development occurs, whether continuous or categorical, in infants’ understanding of the meaning of eye-gaze across this time period.

The difference between the behavioural indices used to ‘measure’ gaze following and responding to JA can be subtle. Unlike gaze following, which is usually measured simply by correct orienting, responding to JA often includes shifting attention back and forth between

the person and the referred object, rather than simple orienting (Carpenter, Nagell, & Tomasello, 1998). Furthermore, looking time to the object is taken as a measure of infants' referential understanding. Brooks and Meltzoff (2005) found that from 10 months infants looked longer at a target object when the adult looked at the object with their eyes open (versus eyes closed). They argue that from at least this age, infants understand the importance of open eyes as a cue to the other person 'seeing' something, and that the adult's looking behaviour causes the object to acquire a new meaning for these infants. In other words it is not just the act of orienting but the subsequent looking behaviour which distinguishes infants who understand the *meaning* of gaze.

3.2.1 Prospective studies of infants at high risk for ASD

Across different studies, impairments in JA behaviours characterise very young children with ASD (for a review see Elsabbagh & Johnson, 2007) and Charman (2003) argues that JA difficulties are likely to be the best 'discriminators' of emerging ASD symptoms in infants between 12 and 18 months. In order to understand the emergence of such JA difficulties in children who later go on to develop ASD, it is necessary to investigate the way in which precursors to JA might influence the trajectory of development. Both gaze following and attentional engagement with the gazed-at object are theoretically important for the development of JA. Considering these behaviours in the context of development is crucial, as the ability to discriminate direction of eye-gaze (Caron, Caron, Roberts & Brooks, 1997) and to flexibly shift attention (Hood & Atkinson, 1993) emerge very early in typical development. Given that a diagnosis of ASD rarely occurs before 2 years of age, looking prospectively at gaze following behaviour in infants at high risk for an ASD is therefore important in understanding the developmental trajectory of this behaviour.

Several prospective studies have investigated 'JA behaviours' in infants at high risk for ASD, although the measures used are in fact often more akin to gaze following behaviour

than JA. Presmanes et al. (2007) tested 12- to 23-month-olds using ten JA prompts in a variety of different combinations. Trials ranged from a single cue (e.g., silent gaze shift) to highly redundant (e.g., gaze shift with point and vocalisation). At these two extremes, the performance of high-risk and control children was similar. However, group differences were found for the intermediate cue conditions (e.g., gaze shift and vocalisation), with reduced looking to the target by high-risk children. In an extension of the study, Yoder, Stone, Walden and Malesa (2009) employed growth curve modelling to examine the relationship between early responding to JA and later social impairment, measured by observation of response to JA at outcome and ASD diagnosis. They found that initial level of response to JA (mean age 15 months), but not its growth rate, predicted response to JA impairment and ASD diagnosis at outcome (34 months). However, as the infants in this study were already 15 months of age at the initial assessment, there is no way of telling whether such early differences were present from the first few months, or alternatively developed during the first year of life.

JA has also been examined prospectively using the Early Social Communication Scales (ESCS; Mundy, Hogan & Doehring, 1996), a standardised clinical observation of JA behaviours. Cassel et al. (2007) demonstrated reduced response to JA on the ESCS in a group of 18-month-old high-risk siblings compared to low-risk children with a typically developing older sibling. However, unlike Yoder et al. (2009), Cassel et al. (2007) found no significant response to JA impairment in the high-risk group at the younger age of 15 months, nor earlier at 8, 10 or 12 months. Other studies using the ESCS in high-risk and low-risk infants younger than 18 months have also shown no significant group differences in responding to JA (e.g. Goldberg et al., 2005; Yirmiya et al., 2006). Negative findings might be attributable to differences in cue-type, with multiple cues in the ESCS (calling the child's name, pointing

and looking at an object) compared to ‘intermediate’ levels of cues in Presmanes et al.’s (2007) experimental paradigm.

More recently, the neural underpinnings of eye-gaze processing in high-risk infants have been examined. Elsabbagh et al. (2009b) used an ERP technique to look at brain responses to direct as compared to averted gaze in high-risk and low-risk 10-month-old siblings. They found that although the two groups could not be differentiated in their response to direct versus averted gaze in early latency posterior ERP components, the high-risk group showed a longer latency in response to direct gaze in a later component. Whilst understanding of communicative intent is not likely to emerge until later in development, such early differences in later components may relate to subsequent JA difficulties in these infants.

Taken together, the studies of responding to JA in high-risk infants suggest that impairments emerge, rather than being present from the earliest point, towards the end of the first year of life. It has previously been suggested that a range of subtle deficits early in development may interact with each other and the environment and become more pronounced over time (Elsabbagh & Johnson, 2010). Thus, studying precursors of JA, including gaze following and looking time to the referred object, may elucidate some of the inconsistent findings described earlier.

3.2.2 The gaze following task

In the current study eye-tracking was used to look at gaze following behaviour in an experimental task (see Senju & Csibra, 2008). In Senju and Csibra’s (2008) task, 6-month-old typically developing infants viewed short videos of a model turning to look at one of two objects. They found that the number of first looks and looking time to the congruent object was significantly greater than that to the incongruent object when the model engaged the watching infant in eye contact before shifting their gaze. A group of infants at high risk for an ASD and low-risk controls from the British Autism Study of Infant Siblings (BASIS) were

tested. Infants took part in the study at around 7 months, and again at around 13 months of age. The high-risk infants were split into TD-sibs, AT-sibs and ASD-sibs on the basis of clinical assessment at 36 months. I aimed to test whether there were group differences between high-risk and low-risk infants, and/or group differences between the ASD-sibs and other infants, in both gaze following behaviour and attentional engagement with the congruent object. Further, I aimed to test whether any such differences were apparent at 7 or 13 months.

3.3 Method

3.3.1 Participants

As discussed in Chapter 2, the current study forms part of a battery of studies administered to infants as part of the *BASIS* study: www.basisnetwork.org.uk. Out of the total sample, only the 73 infants (35 high-risk and 38 low-risk) who completed the gaze following task at both visits (7 and 13 months) were included in the analysis (mean time between visits: low-risk mean = 6.45 months, standard deviation; $sd = 0.98$; high-risk mean = 6.37 months, $sd = 0.91$). Independent samples t-tests showed the groups did not differ significantly on age at either visit: 7 month visit (low-risk mean = 7.4, $sd = 0.11$, range 6-10; high-risk mean = 7.1, $sd = 0.09$, range 6-10), $t(71) = 0.709$, $p = 0.481$; 13 month visit (low-risk mean = 13.8, $sd = 0.10$, range 11-16; high-risk mean = 13.6, $sd = 0.13$, range 11-18), $t(71) = 0.823$, $p = 0.413$.

Table 3.1 Descriptive Statistics for the Mullen Scales of Early Learning (*MSEL*) Early Learning Composite Standard Scores and ADOS-G Scores

	Low-risk controls		High-risk infants		'TD-sibs'		'AT-sibs'		'ASD-sibs'	
	M	S.E	M	S.E	M	S.E	M	S.E	M	S.E
7m <i>MSEL</i> ELC	n = 38		n = 35		n = 14		n = 9		n = 12	
	105.74	1.91	91.49	2.33	92.43	2.94	89.89	2.39	91.58	5.78
13m <i>MSEL</i> ELC	n = 38		n = 35		n = 14		n = 9		n = 12	
	108.00	2.52	97.23	2.39	99.93	3.15	99.78	3.41	92.17	5.29
24m <i>MSEL</i> ELC	n = 33		n = 33		n = 14		n = 8		n = 11	
	116.52	2.53	102.85	3.37	109.57	3.77	96.50	5.04	98.91	7.88
36m <i>MSEL</i> ELC	n = 37		n = 32		n = 13		n = 9		n = 10	
	116.08	2.66	105.03	3.84	113.77	3.02	100.44	6.41	97.80	9.73
24m ADOS-G score			n = 34		n = 14		n = 9		n = 11	
			6.88	0.66	4.64	0.77	6.33	0.76	10.18	1.15
36m ADOS-G score	n = 38		n = 35		n = 14		n = 9		n = 12	
	5.05	0.72	8.80	0.96	4.00	0.57	11.33	1.44	12.5	1.62

3.3.2 Behavioural assessment and outcome groups

Infants were assessed on the *Mullen Scales of Early Learning* (*MSEL*; Mullen, 1995) at 7, 13, 24 and 36 months (see Table 3.1). The low-risk control and high-risk groups showed significantly different early learning composite scores at both visits: 7 month visit, $t(71) = 4.77$, $p < 0.001$; 13 month visit, $t(71) = 3.09$, $p = 0.003$. At the 24 month visit the *Autism Diagnostic Observation Schedule – Generic* (ADOS-G; Lord et al., 2000) assessment was administered to high-risk children only. At 24 months, of the 35 children who took part in the

gaze following task at both the 7 and 13 month visits, two toddlers completed Module 2 and 32 completed Module 1. One child did not take part in the 24 month visit but was still included in group analysis of gaze following data. At 36 months, both groups completed the ADOS-G assessment. Of the 35 high-risk children, 33 completed Module 2 and two toddlers completed Module 1. All 38 low-risk control children completed Module 2. Assessments were administered by trained researchers who had not previously seen the children at the 7 month or 13 month visit, and were thus blind to infant performance on experimental measures. All ADOS-G assessments were double coded and a consensus code was agreed by the researchers. Intra-class correlation (ICC) coefficients between coders was very high (24 months ICC = 0.73; 36 months high-risk ICC = 0.76, low-risk ICC = 0.87).

Following the classification procedure in Chapter 2, toddlers from the high-risk group were split into three outcome groups: typically developing (TD-sibs) atypically developing (AT-sibs) and ASD (ASD-sibs). From the 35 high-risk infants in the current study seen for diagnostic assessment at 36 months, 12 (8 boys) met criteria for an ASD diagnosis (34.3%), 14 (4 boys) were in the TD-sibs group (40%) and 9 (2 boys) were in the AT-sibs group (25.7%). Of those classified as AT-sibs, 6 scored above ADOS-G cut-off for ASD, 1 scored above ADOS-G cut-off for ASD and below *MSEL* 1.5 SD cut-off, 1 above ADI-R cut-off, and 1 scored below *MSEL* 1.5 SD cut-off.

3.3.3 Apparatus

Infants' looking behaviour was recorded using a Tobii 1750 eye-tracker. The eye-tracker has an infrared light source and a camera mounted below a 17-inch flat screen monitor to record corneal reflection data. To evaluate where on the screen the infant is looking, the Tobii system used measurements of gaze direction from each eye separately. Stimuli were presented on the screen using ClearView software. Infants sat on their parent's lap 50 centimetres away from the screen. The distance and height of the screen were adjusted for

each infant in order to get good tracking of their eyes. Before starting the main experimental task, a five-point calibration sequence was run. The eye-tracking task was started when at least 4 points were marked as correctly calibrated for each eye. Gaze data were recorded at 50 Hz, and the spatial resolution was 1° after calibration.

3.3.4 Stimuli and procedure

Stimuli used in this study were the same as those used in Senju and Csibra (2008). Example stills from trials presented to the infant are displayed in Figure 3.1. Each sequence began with two objects on a table and a female model ‘looking down’ (3 seconds), then looking up – ‘direct gaze’ (2 seconds) – and then turning her head to look at one of the objects – ‘shift’ (6 seconds). The ‘looking down’ phase was measured from the start of the trial until the model looked up and both her head and eye-gaze were directed straight ahead. The ‘direct gaze’ phase began as soon as the model’s eyes were looking ahead, and finished when her head began to turn away. This turning marked the beginning of the ‘shift’ phase, which finished at the end of the trial. The object looked at by the model during ‘shift’ is the *congruent* object, and the other, non-gazed at object is the *incongruent* object. Each infant viewed 12 trials, and there were six different pairs of objects whose position with respect to the gaze was counterbalanced across trials. Thus in different trials the same object would once be the congruent object and once the incongruent object. The direction of the female’s gaze shift was fixed in the following pseudo-random order: RLLRLRRLRLR. Before the beginning of each trial, the infants’ attention was directed to the screen using small animations.



Figure 3.1 Screen shots of the videos presented to infants, split into the three phases used in analysis: looking down, direct gaze, shift; and with the congruent and incongruent object areas of interest (AOIs) highlighted. The visual angle of the overall screen took up 37.6° horizontally and 30.5° vertically. Depending on their size, the visual angle of the objects, subtended 3.7° by 4.5° for the smallest and 7.3° by 8.4° for the largest.

3.3.5 Data analysis

For the purpose of trial exclusion, all trials were split temporally into three phases: looking down, direct gaze and shift (see Figure 3.1). Within each trial, three rectangular areas of interest (AOIs) were defined around the face, congruent object and incongruent object using ClearView software (face subtended 8° by 11.4° and objects by 3.7° by 4.5° for the smallest and 7.3° by 8.4° for the largest). Gaze data were extracted for each of these AOIs as well as a total for the whole slide, using a fixation filter of 60 m/s to exclude random noise unlikely to represent true fixations. Trial exclusion criteria were: (i) no looking to the face during ‘direct gaze’ as Senju and Csibra (2008) found this to be a prerequisite for gaze following behaviour; and (ii) looking away from the computer screen for the entire ‘shift’ phase. Only data from the final ‘shift’ phase was used to calculate the measures of interest: gaze following, defined as a higher proportion of first looks to the congruent than the incongruent object, and attentional engagement, defined as looking time to congruent object for all trials in which the first look was correct.

First look. For analysis of first look responses, infants who completed < 3 valid trials were excluded. The number of valid trials for first look did not differ across groups at either visit (see Table 3.2; 7 month visit: $F(2, 70) = 1.406$, $p = 0.252$; 13 month visit: $F(2, 70) = 0.539$, $p = 0.586$). First look responses were measured from the beginning of the ‘shift’ phase, and calculated for the congruent and incongruent objects. In this study *proportion of first looks to the congruent versus incongruent object* (Moore & Corkum, 1998) was chosen as the primary measure of gaze following behaviour. This measure reflects infants’ ability to follow the direction of another person’s gaze to its target. Trials in which the infant did not orient to either object, but was still looking at the screen were included when calculating proportions. These ‘other’ trials included infants being stuck on the face or orienting to other parts of the screen and they contributed to the denominator of the number of trials and were treated as the reference category in model parameters defining proportions. See Appendix 5, Figure A5.1, for the distribution of first look difference scores (i.e., proportion of congruent – incongruent trials).

Looking time. Following Brooks and Meltzoff (2005) and Senju and Csibra (2008), I also chose to analyse looking time behaviour, as this is thought to reflect referential understanding as well as being robust against noise arising from any brief loss of tracking. Looking time behaviour was analysed only for trials in which infants were correct in their first look, and a further 4 infants were excluded as they had no correct first looks. Of these excluded infants, 3 were from the high-risk group, and they did not show *MSEL* or *ADOS-G* scores systematically higher or lower than the other infants. There were no significant group differences in total looking time at either visit (see Table 3.2, 7 month visit: $F(2, 66) = 1.69$, $p = 0.192$; 13 month visit: $F(2, 66) = 1.148$, $p = 0.324$). Attentional engagement was defined as looking time to the congruent object (out of total looking time to the slide) during the ‘shift’ phase, for all first look trials that were correct. This measure reflects not only the infants’

ability to follow gaze but also their subsequent engagement with the target of another person's gaze. Looking time to 'other' parts of the screen (face, torso, blank sides and table) was included in the denominator as the reference category for parameters defining proportions. See Appendix 5, Figure A5.2, for looking time distributions.

Table 3.2 Number of Valid Trials and Total Duration of Looking (ms) at the 7 and 13 Month Visits by Group and Condition

	Low-risk controls		High-risk infants		‘TD-sibs’		‘AT-sibs’		‘ASD-sibs’	
	M	S.E	M	S.E	M	S.E	M	S.E	M	S.E
First look 7m	n = 38		n = 35		n = 14		n = 9		n = 12	
Valid trials	8.55	0.41	8.60	0.51	8.29	0.85	8.56	0.92	9.00	0.93
Attentional	n = 37		n = 32		n = 14		n = 7		n = 11	
Engagement 7m										
Valid trials	3.24	0.30	2.84	0.25	2.71	0.35	3.14	0.63	2.82	0.42
Total duration	3792	202	3243	246	3216	295	2967	565	3452	519
First look 13m	n = 38		n = 35		n = 14		n = 9		n = 12	
Valid trials	9.47	0.34	9.86	0.32	9.93	0.37	9.89	0.59	9.75	0.72
Attentional										
Engagement 13m	n = 37		n = 32		n = 14		n = 7		n = 11	
Valid trials	5.51	0.35	5.44	0.42	5.43	0.61	5.14	1.20	5.64	0.62
Total duration	4231	156	4386	145	4239	234	4582	351	4448	218

3.3.6 Generalised estimating equation (GEE) analysis

These repeated measures multinomial data were analysed as a set of simple correlated proportions using a generalised estimating equation approach (Pickles, 1998). The GEE approach allows the first look data to be treated as binomial (as responses were either correct

or incorrect) and the looking time data as normally distributed. In the first look analyses, for each of the two assessments the number of responses in each category were analysed as a count, with the total number of each infant's responses at that assessment occasion as a binomial denominator and with a logit link between predictors and the expected proportion. In the looking time analyses, for each of the two assessments the proportion of time in each category was analysed as a Gaussian model with identity link between predictors and the expected proportion. Another advantage of the GEE method comes from the fact that in all cases Wald tests were used to determine the significance of effects, calculated from the sandwich estimator of the parameter covariance matrix. These tests are therefore robust to errors in the assumed correlation between the response proportions and also to the variation in the precision or overdispersion in proportions arising from the varying amounts of valid trials available from each infant.

3.4 Results

For each visit (7 and 13 months), the control and high-risk groups were first compared on experimental measures (first look and looking time), to examine overall group difference based on risk status. The relationship with clinical outcome was then examined, with the high-risk infants split into three groups: 'TD-sibs', 'AT-sibs', 'ASD-sibs'. For all analyses, the *MSEL* composite standard score (from either 7 month or 13 month visit) was included as a covariate, to account for any group differences in IQ.

3.4.1 Seven month visit

For the 7-month-old infants (see Table 3.3), a generalised estimating equation showed that both low-risk controls and high-risk infants followed gaze, looking significantly more to the congruent than incongruent object (coef = 0.50, S.E = 0.14, $z = 3.7$, $p < 0.001$), with no significant difference between groups (coef = -0.01, S.E = 0.33, $z = 0.03$, $p = 0.98$).

Proportion of looking time was then calculated for all *correct* first look trials. No group differences between high-risk and low-risk infants were found for looking to the congruent object (coef = 0.03, S.E = 0.05, $z = 0.52$, $p = 0.61$) (see Table 3.3).

Table 3.3 Proportion of First Looks to and Attentional Engagement with the Congruent and Incongruent Objects at the 7 Month Visit

	Low-risk controls		High-risk infants		'TD-sibs'		'AT-sibs'		'ASD-sibs'	
	M	S.E	M	S.E	M	S.E	M	S.E	M	S.E
First look	n = 38		n = 35		n = 14		n = 9		n = 12	
Congruent object	0.40	0.04	0.34	0.03	0.31	0.04	0.36	0.04	0.37	0.05
First look	n = 38		n = 35		n = 14		n = 9		n = 12	
Incongruent object	0.28	0.03	0.23	0.03	0.24	0.05	0.25	0.07	0.20	0.05
Attentional Engagement	n = 37		n = 32		n = 14		n = 7		n = 11	
Congruent object	0.28	0.03	0.31	0.04	0.31	0.06	0.27	0.05	0.34	0.08
Attentional Engagement	n = 37		n = 32		n = 14		n = 7		n = 11	
Incongruent object	0.03	0.01	0.04	0.01	0.03	0.01	0.04	0.03	0.05	0.03

High-risk infants were then split into TD-sibs, AT-sibs and ASD-sibs based on clinical outcomes at 36 months. No significant group difference in the proportion of first looks to the congruent versus incongruent object were found (coef = 0.04, S.E = 0.15, $z = 0.28$, $p = 0.77$). Nor were there any significant group differences in terms of looking time to the congruent object ($\chi^2(3) = 1.1$, $p = 0.78$).

3.4.2 Thirteen month visit

At 13 months (see Table 3.4), low-risk controls and high-risk infants both had a significantly higher proportion of first looks to the congruent than incongruent object (coef = 1.27, S.E = 0.16, $z = 8.06$, $p < 0.001$), with no group interaction (coef = -0.008, S.E = 0.33, $z = 0.02$, $p = 0.98$). Nor were there any significant group differences between high-risk and low-risk infants in looking time to the congruent object (coef = -0.05, S.E = 0.03, $z = 1.57$, $p = 0.12$) (see Table 3.4).

Table 3.4 Proportion of First Looks to and Attentional Engagement with the Congruent and Incongruent Objects at the 13 Month Visit

	Low-risk controls		High-risk infants		'TD-sibs'		'AT-sibs'		'ASD-sibs'	
	M	S.E	M	S.E	M	S.E	M	S.E	M	S.E
First look	n = 38		n = 35		n = 14		n = 9		n = 12	
Congruent object	0.57	0.03	0.51	0.04	0.54	0.05	0.46	0.10	0.51	0.07
First look	n = 38		n = 35		n = 14		n = 9		n = 12	
Incongruent object	0.28	0.02	0.25	0.03	0.26	0.05	0.21	0.05	0.26	0.05
Attentional Engagement	n = 37		n = 32		n = 14		n = 7		n = 11	
Congruent object	0.31	0.02	0.26	0.02	0.30	0.03	0.21	0.02	0.22	0.03
Attentional Engagement	n = 37		n = 32		n = 14		n = 7		n = 11	
Incongruent object	0.08	0.01	0.06	0.01	0.08	0.02	0.03	0.01	0.06	0.02

When split by outcome group, no significant interaction with group was found for first looks to the congruent versus incongruent object (coef = -0.02, S.E = 0.15, $z = 0.1$, $p = 0.91$). For looking time, a significant overall group interaction was found for looking to the congruent object ($\chi^2(3) = 13.11$ $p = 0.004$; see Figure 3.2) with significantly reduced looking

time in the AT-sibs compared to controls (coef = -0.09, S.E = 0.03, $z = 2.88$, $p = 0.004$) and TD-sibs (coef = -0.09, S.E = 0.03, $z = 2.72$, $p = 0.007$), and significantly reduced looking time in the ASD-sibs compared to controls (coef = -0.08, S.E = 0.04, $z = 2.21$, $p = 0.03$) and TD-sibs (coef = -0.08, S.E = 0.04, $z = 2.14$, $p = 0.03$). There were no significant differences either between TD-sibs and controls (coef = -0.003, S.E = 0.02, $z = 0.1$, $p = 0.92$) or between AT-sibs and ASD-sibs (coef = 0.01, S.E = 0.03, $z = 0.44$, $p = 0.66$).

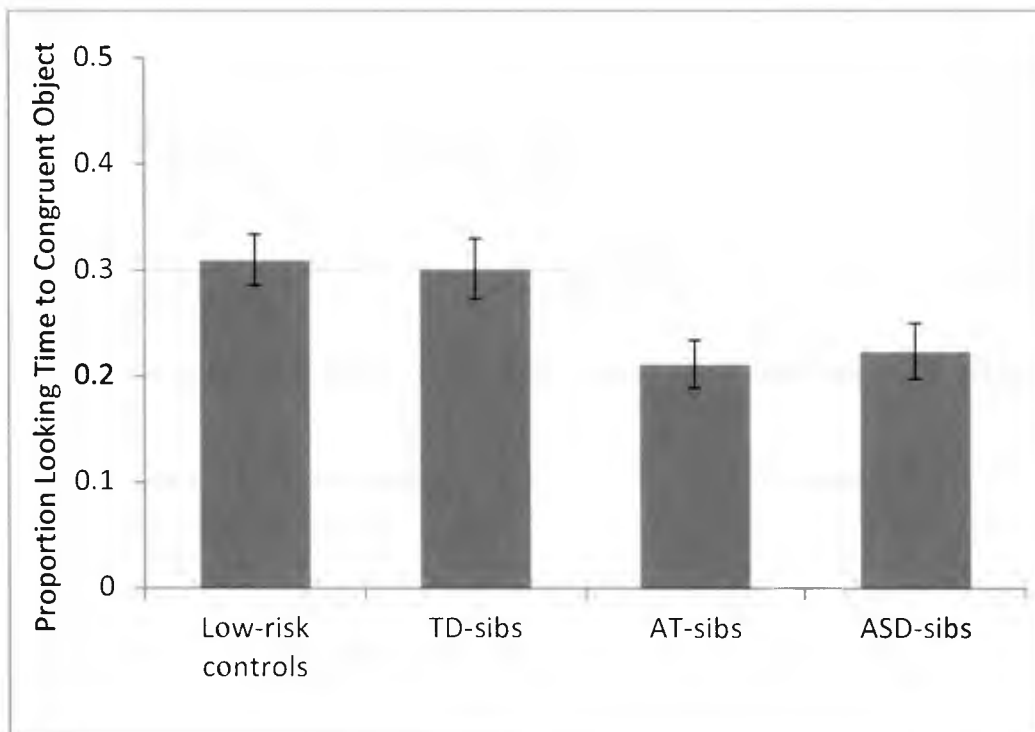


Figure 3.2 Attentional engagement: Proportion of looking time to the congruent object at the 13 month visit. Error bars +/- 1 S.E.

Given the finding of reduced looking time to the congruent object in both AT-sibs and ASD-sibs, it is possible that this effect is associated with autistic-like social communication difficulties in general rather than a categorical clinical diagnosis, as both groups show high scores on the 24 and 36 month ADOS-G (see Table 3.1). Within the high-risk group, a partial correlation accounting for *MSEL* composite score at the 13 month visit showed a significant

negative correlation between looking time to the congruent object at 13 months and continuous 24 month ADOS-G score ($r = -0.46$, $p = 0.01$). At 36 months, both controls and high-risk infants were assessed on the ADOS-G. The overall correlation with 13 month looking time to the congruent object did not reach significance although the trend was in the same direction ($r = -0.22$, $p = 0.07$). When split by group, it is the high-risk infants ($r = -0.31$, $p = 0.09$) driving this effect, rather than the low-risk controls ($r = -0.1$, $p = 0.56$).

3.5 Discussion

This study aimed to determine whether early problems in spontaneous gaze following and looking behaviour during infancy is part of the broader autism phenotype (i.e., low-risk versus high-risk group differences), or whether such difficulties relate to ASD outcome at 3 years. For the controls, if the stringent definition of gaze following is adopted, of a higher proportion of first looks to the congruent than incongruent object (Moore & Corkum, 1998), then like Senju and Csibra (2008) both 7- and 13-month-old infants were able follow gaze. Our finding that gaze following was neither influenced by risk status or by later emerging social and communication difficulties measured by the ADOS-G and categorical ASD outcome status within those at high risk, suggests that the early mechanisms for automatic orienting to another's gaze are intact. This is unsurprising in that such orienting is present in other primates (Tomasello, Call & Hare, 1998) whereas JA arguably is uniquely human (Povinelli, Bierschwale & Cech, 1999; Baron-Cohen, 1995). However, by 13 months, reduced looking to the congruent object after gaze following was found in high-risk infants who go on to an ASD or atypically developing outcome at 3 years. This suggests that having followed gaze direction, these infants may not use this information to preferentially attend to the gazed-at object.

It was expected that low-risk control infants would show gaze following behaviour, indexed by significantly more first looks to the congruent than the incongruent object, at both

visits. The results supported this, with controls showing a greater proportion of first looks to the congruent than the incongruent object at both 7 and 13 months. However, it was also hypothesised that group differences between controls and high-risk infants would either be present from the 7 month visit and persist over time, or emerge later, at the 13 month visit. Contrary to our hypothesis, no group differences were found, with the high-risk infants also following gaze at both visits. Intact orienting to the congruent versus incongruent object was also found for infants later classified as having ASD.

This finding of gaze following behaviour in high-risk infants, even those who go on to develop ASD, might seem surprising given the evidence that real-life difficulties in responding to another's gaze are one of the earliest discriminators of children who go on to develop ASD (Charman, 2003). However, according to a review by Nation and Penny (2008), unlike in 'real-life settings' the majority of published papers find no evidence for deficits in orienting to social stimuli in children already diagnosed with ASD when an experimental design is utilised. For example, Swettenham, Condie, Campbell, Milne and Coleman (2003) found that, like typically developing children and adults (e.g., Hood, Willen & Driver, 1998; Driver et al., 1999), children with ASD are also faster to orient to objects cued by moving eye-gaze. Senju, Tojo, Dairoku and Hasegawa (2004) demonstrated the existence of this cueing effect in high-functioning autistic children and controls even when it was made explicit that the cue was counter-informative (on 80% of the trials the object appeared in the uncued location). Chawarska, Klin and Volkmar (2003) used both a Posner experimental task and the ADOS-G to examine JA deficits in 2-year-old children with ASD. They found that whilst deficits in JA were pronounced in the ADOS-G assessment, a cueing effect of eye-gaze was nevertheless observed in the experimental measure. These findings are in line with our results that early automatic orienting to another's gaze is not impaired in the broader autism phenotype in infancy, nor is it a predictor of ASD outcome.

For all trials in which first look was correct, a measure of looking time to congruent object was calculated. Attentional engagement with the target of another person's gaze is a measure more associated with referential understanding of the gaze (Brooks & Meltzoff, 2005). No statistically significant group differences between high-risk and control infants were found at either visit. When split by outcome, looking time to the congruent object at 13 months was significantly reduced in the ASD-sibs and AT-sibs compared to controls *and* TD-sibs. Reduced looking to the congruent object at 13 months suggests that whilst these infants who go on to show ASD or atypical development (i.e., autistic-like social communication difficulties and/or language or developmental delays) are able to orient correctly in response to the gaze shift, they may not be sensitive to the referential nature of the gaze. This finding provides empirical evidence for Sullivan et al.'s (2007) observation that children with ASD who were able to pass a JA task 'did not seem to understand the intent of the examiner'. Further, they noted 'an "emptiness" of gaze' and some children locating the target but then quickly looking away to a different object.

Given the mixed picture of impairment and non-impairment depending on the measure chosen (first look versus looking time) it is useful to understand which of these behaviours is consequential for learning or for developing typical social interactions. As discussed, Brooks and Meltzoff's (2002; 2005) studies suggest that looking behaviour distinguishes infants who understand the meaning of eye-gaze. This is in line with an ERP study in 9-month-olds, in which infants saw another person looking at an object, either preceded by a period of mutual gaze (JA) or not (non-JA) (Striano, Reid & Hoehl, 2006). Infants showed an increased amplitude of a neural correlate reflecting attentional engagement when processing an object looked at by another person, as compared to an object in a non-JA situation. Taken together these studies suggest that the gaze of another person can influence subsequent object processing in infants, at both a neural and behavioural level.

Looking time at 13 months distinguishes not only infants who go on to develop ASD, but also those who show atypical development as measured by the ADOS-G, ADI or *MSEL*. Both these groups show high levels of social communication difficulties, as measured by the ADOS-G. I therefore examined the correlation between continuous ADOS-G score and looking time to the congruent object at 13 months and found a significant relationship within the high-risk infants with the 24 month ADOS-G. There was also a non-significant trend with the later 36 month ADOS-G in the high-risk, but not control infants. This suggests that our looking time measure relates to social communication behaviour in high-risk infants, rather than ASD outcome per se. The fact that this correlation is weaker with the 36 month ADOS-G is probably due to the increased time between measurement occasions.

The result of no group difference at the early 7 month visit is consistent with findings from other high-risk sibling studies looking for behavioural markers within the first year of life (e.g., Elsabbagh & Johnson, 2010; Rogers, 2009; Yirmiya & Charman, 2010). More specifically to our study, the lack of an outcome group difference in attentional engagement at the 7 month visit could be due to the development, over the period 7–13 months, of an understanding of the meaning of gaze, and this is impaired in the infants with social communication difficulties. This fits with Tomasello et al.'s (2005) model in which they argue that infants develop from being able to follow the direction of gaze at 6 months to a full understanding of intentionality around 14 months. Alternatively, group differences in looking behaviour may have been present earlier in development but not measured by our task, either because the task was not sufficiently sensitive to detect such differences, or the differences were too subtle to measure behaviourally. It is possible that there were early group differences in neural processing (Elsabbagh et al., 2009b; Elsabbagh et al., 2012) which compounded over time contributed to the emergence of reduced looking time by 13 months.

Understanding the precise similarities and differences between the measures used in this task as compared to previous studies is important when thinking about replication of the current findings. This task is based on one condition from Senju and Csibra's (2008) gaze following study in 6 month old infants. They calculated difference score measures for first look, looking time and frequency of shifts to the congruent versus incongruent objects across four conditions, eye contact (the same as this task), non-eye contact (face masked by a cartoon), infant directed speech and adult directed speech. The first look measure in this task is broadly comparable to first look in their eye-contact condition, although here trials with first looks to other parts of the screen were included in the analysis as the denominator, rather than discarded as invalid. The results for first look are similar across studies with a higher proportion to the congruent than the incongruent object.

For looking time it was not really possible to compare directly across tasks, as looking time here was analysed only for trials in which the first look was correct (not the congruent versus incongruent object difference score used by Senju and Csibra, 2008). However, my measure is similar to the looking time measure used by Brooks and Meltzoff (2002), in which they calculated the total duration of correct looks divided by the number of trials with a correct look. Despite the similarity between the looking time measures it is not really possible to directly compare across methodologies, with looking time in this study measured by an eye-tracker as compared to video coding used by Brooks and Meltzoff (2002).

Thus, while the measures in this task are broadly similar to those used by other studies, there are differences in the treatment of valid trials and in the methodology used to record gaze behaviour, which limit the parallels which can be drawn. Future work will be important to replicate these results.

There is clear evidence that difficulties in responding to JA characterise young children with ASD. Given the links between JA and subsequent socio-communicative development, a

key area of impairment in ASD, it is plausible that such behaviours play an etiological role in the condition. To understand the developmental pathways leading to diagnosis it is necessary to look at precursors to JA, including gaze following behaviour. This experimental task used eye-tracking to derive measures of gaze following. Whilst not as ecologically valid as a response to JA task in a naturalistic environment, this paradigm can be used in much younger infants. It is also possible to calculate different measures related to gaze following accurately, such as subsequent attentional engagement with the target object. Further, although it was a computer-based task, the stimuli were dynamic video clips of a model turning to look at an object, and thus more ecologically valid than some attention cueing paradigms. Future research should combine this task with ERP methods over a wider developmental time-frame in order to establish whether differences in neural processing precede behavioural differences in looking time responses. While in this study there were no group differences in the number valid trials, it would be interesting in future work to look at whether later gaze following difficulties emerge because children who go on to develop ASD reduce their orienting towards faces and therefore miss the referential cues (see Vivanti et al., 2011).

In conclusion, it was found that gaze following at 7 and 13 months was not impaired in infants at risk for ASD, neither in those who went on to show subsequent sub-threshold social communication or other developmental difficulties at 36 months nor in infants who were classified as having an ASD. However, having followed gaze correctly, infants with later social-communication problems, both those with ASD and atypical development, showed a reduction in looking time to the congruent object by the 13 month visit. This reduced attention may reflect difficulties in understanding the communicative relevance of eye-gaze and be part of the on-going developmental process that leads to an ASD presentation and other developmental atypicalities.

Part 2: Prediction of 24 Month Parent-Reported Vocabulary by Early Gaze Following Behaviour

3.6 Gaze following and vocabulary

As well as predicting atypical outcomes, it is also interesting to see whether individual differences in gaze following behaviours relate to later typical outcomes. As discussed in Chapter 1, predicting outcomes from infancy measures has been an area of research interest for decades. One reliable and widely replicated finding is that early measures of JA predict concurrent and later language comprehension. This is found in both typical development (e.g., Carpenter et al., 1998; Morales et al., 2000) and in children with ASD (e.g., Charman et al., 2003; Siller & Sigman, 2008). This relationship with language development makes sense from a theoretical point of view because gaze cues offer a potential strategy for identifying the referent of a heard word, thus enabling the formation of word-object mappings (e.g., Baldwin, 1995).

3.6.1 Links between JA and vocabulary in typical development

As discussed in Part 1, in typical development infants begin to follow gaze during the first 12 months (Scaife and Bruner, 1975). Thus sharing attention with another person precedes linguistic development. Mundy and Newell (2007) argue that this finding resulted in a reconceptualization in thinking from Piaget's 'egocentric' infant, as it suggests that shared reference with another is possible before the onset of language.

Perhaps the first study to comprehensively examine the longitudinal link between early JA and language development was conducted by Carpenter et al. (1998). They assessed the amount of time infants aged 9-15 months spent in joint engagement with a caregiver together with infants' expressive and receptive vocabulary measured by the CDI at 9-24 months. Carpenter et al. (1998) found that what they termed 'joint engagement', the amount of time

spent in joint attention with the caregiver, at 11-13 months of age was related to concurrent and subsequent language comprehension, but not expression. Morales et al. (2000), on the other hand, found that individual differences in JA at 6, 8, 10 and 18 months of age were related to both receptive and expressive language development at 30 months, and a regression analysis demonstrated this association to be significant using an aggregate measure across 6-18 months, even when 24 month vocabulary was controlled for.

Baker and Nelson (1984) view JA in typical development as providing a ‘scaffold’ for emerging communication and Mundy and Newell (2007) further argue that JA facilitates social learning and development. Understanding this relationship in ASD, which is characterised by language difficulties, may provide insights into possible JA-mediated language intervention strategies.

3.6.2 Links between JA and vocabulary in ASD

As discussed earlier, JA behaviours are among the earliest predictors of later ASD. There is also evidence that, as in typical development, JA predicts both concurrent and longitudinal language abilities in autistic children. Dawson et al. (2004) used path analysis to show a link between JA and concurrent language abilities in 3- to 4-year-old children with ASD. They showed correlations with both receptive and expressive language, for a standardised assessment (*MSEL*) and a parent report measure (Vineland Adaptive Behaviour Scales; *VABS*). Sigman and Ruskin (1999) demonstrated a link between responding to JA bids (mean age 47 months) and long-term gains in expressive language skills nearly nine years later. Further, Charman et al. (2003) demonstrated that JA (measured by gaze switching between a toy and the experimenter or parent) at 20 months predicted 42 month receptive but not expressive language in much younger children with ASD.

More complex growth curve analysis has been employed to look change in language development in children with ASD and Siller and Sigman (2008) found that responding to JA

was a strong predictor of the rate of language growth, even after accounting for variation in IQ and initial language levels. Further, Anderson et al. (2007) showed that children with better JA skills were more likely to develop phrase speech.

3.6.3 Infants at high risk for ASD

The study by Presmanes et al. (2007) discussed in Part 1, as well as looking at different JA cue types, also looked at JA in relation to language in high-risk versus low-risk control infants. They found that response to JA was related to both *MSEL* receptive language (RL) and expressive language (EL) in the high-risk infants but not in the low-risk controls.

Sullivan et al. (2007) also looked prospectively at high-risk infants and showed that response JA at 14 months predicted both later RL and EL as measured by the *MSEL*.

3.6.4 Links between gaze following and language

While there have been various studies looking at the links between JA and language in typical development, far less attention has been focused on the relationship with early gaze following behaviours. Brooks and Meltzoff (2005) conducted a study which looked more specifically at how two different measures of gaze following, ‘looking score’, a measure of gaze following based on correct, incorrect and other first looks and ‘correct first looks accompanied by vocalisation’ at 10-11 months predicted subsequent 14 and 18 month CDI scores, including both receptive and expressive vocabulary. Whilst looking score was not predictive of either measure, correct gaze with simultaneous vocalisation significantly predicted receptive but not expressive language at both 14 and 18 months. This correlation between correct looks with vocalisation and later language comprehension was not accounted for by an overall increased level of vocalisation at 10-11 months, as overall vocalisations did not correlate with later CDI scores.

Brooks and Meltzoff (2005) argue that the combination of simultaneous vocalising with correct looking behaviour reflects a proto-declarative effort by the infant to engage in what

they term ‘psychological sharing’, which is why this measure correlates with later receptive vocabulary. They suggest that it may not predict expressive language because, unlike language comprehension, expression requires articulatory skills (e.g., Bates, 1993).

3.6.5 *The current study*

The looking time and first look measures derived from the gaze following task in Part 1 are different to those chosen by Brooks and Meltzoff (2005). However, the looking score measure in Brooks and Meltzoff’s (2005) study is somewhat similar to the first look measure here, with both reflecting an infant’s ability to follow gaze. Further, although the correct first look with vocalisation is different to the measure chosen here of looking time following a correct first look, both measures are thought to index some level of social understanding.

Following Brooks and Meltzoff (2005), it was hypothesised that irrespective of group membership, a relationship would be found between 13 month ‘looking time’ and both concurrent and subsequent receptive vocabulary. However, no relationship between first look and vocabulary was predicted. I also aimed to extend the Brooks and Meltzoff (2005) study, first, by examining gaze following at a younger age, and testing whether the relationship between looking time and later language extended to the 7 month visit, and second, by controlling for non-verbal IQ score in the relationship between looking time and vocabulary.

3.7 *Method*

3.7.1 *Data analysis*

MacArthur-Bates CDI: A version of the CDI (Fenson et al., 1993) slightly adapted for British English (e.g., ‘trash’ was changed to ‘rubbish’) was completed by parents of infants at 12 and 24 months. At 12 months the *words and gestures* form was used which includes two columns (‘understands’ and ‘understands and says’), thus enabling raw receptive and expressive language scores to be calculated. The *words and sentences* version was used at 24 months.

This form, which normally only includes an ‘understands and says’ column was modified to include an ‘understands’ column, the same as the *words and gestures* form, to allow computation of equivalent expressive and receptive vocabulary counts at both visits.

Although the CDI was completed by parents at all four visits, in this analysis only data from the 13 and 24 month visits are presented, due to extreme floor and ceiling effects at the 7 and 36 months visits respectively.

Gaze following measures: The two gaze following measures used are identical to those presented in Part 1.

1) Difference score of the proportion of first looks to the congruent object – proportion of first looks to the incongruent object. Note that on some trials, first looks were to other parts of the screen and this is reflected in the proportions of congruent and incongruent first looks (which don’t add up to 1) but is not included in the calculation of the difference score.

2) Attentional engagement: proportion of looking time to the congruent object for all first look trials that were correct.

Mullen Scales of Early Learning (MSEL, Mullen, 1995): The fine motor (FM) and visual reception (VR) subscales were also used to create a non-verbal t-score, $(FM+VR)/2$. This was used in partial correlations as a covariate.

3.8 Results

Relationship with the CDI

First, correlations were run on the whole sample between the gaze following measures (looking time and first look) and 13 and 24 month raw CDI scores for receptive and expressive language (see Table 3.5).

Table 3.5 Correlations between Gaze Following Performance and 24 Month Receptive Vocabulary Size

CDI	First look 7m	First look 13m	Attentional Engagement 7m	Attentional Engagement 13m
Receptive 24m				
<i>Overall</i>				
<i>Corr. coef</i>	0.09	-0.12	0.05	0.30*
<i>p</i>	ns	ns	ns	.02
<i>n</i>	65	65	63	63
<i>Low-risk</i>				
<i>Corr. coef</i>	0.08	-0.3	0.13	0.30
<i>p</i>	ns	0.07	ns	0.07
<i>n</i>	37	37	36	36
<i>High-risk</i>				
<i>Corr. coef</i>	0.22	0.05	0.05	0.14
<i>p</i>	ns	ns	ns	ns
<i>n</i>	28	28	27	27

The only significant correlation was between 13 month looking time and subsequent 24 month receptive vocabulary $r = 0.298$, $p = 0.018$. However, examination of the scatterplot in Figure 3.3 showed the presence of an outlier with a very high proportion of gaze to the congruent object (0.91) with a standard residual of -2.22. There is no reason to believe this outlier has arisen due to a data entry error, as the data are extracted from the eye tracker and copied into an excel spread sheet. The outlier was removed to re-test the correlation, which became non-significant although the trend remained the same ($r = 0.22$, $p = 0.086$) suggesting that this single individual point is disproportionately driving the relationship between 13

month looking time and CDI receptive language at 24 months. However, given the limited sample size, and the fact that the outlier is less than three standard deviations from the mean, rather than removing this value completely, it may be better to replace it with a value one point above the highest non-outlier score (i.e., 0.61, see Tabachnick & Fidell, 2007).

Following this procedure gives a significant correlation, $r = 0.28$, $p = 0.03$.

It is clear that, while how we treat this point influences the level of significance of the correlation, whichever method we apply the trend is still in the direction of increased looking time to the congruent object at 13 months predicting higher 24 month receptive language abilities. Thus, for all the following analyses, the outlier was replaced with a value of 0.61.

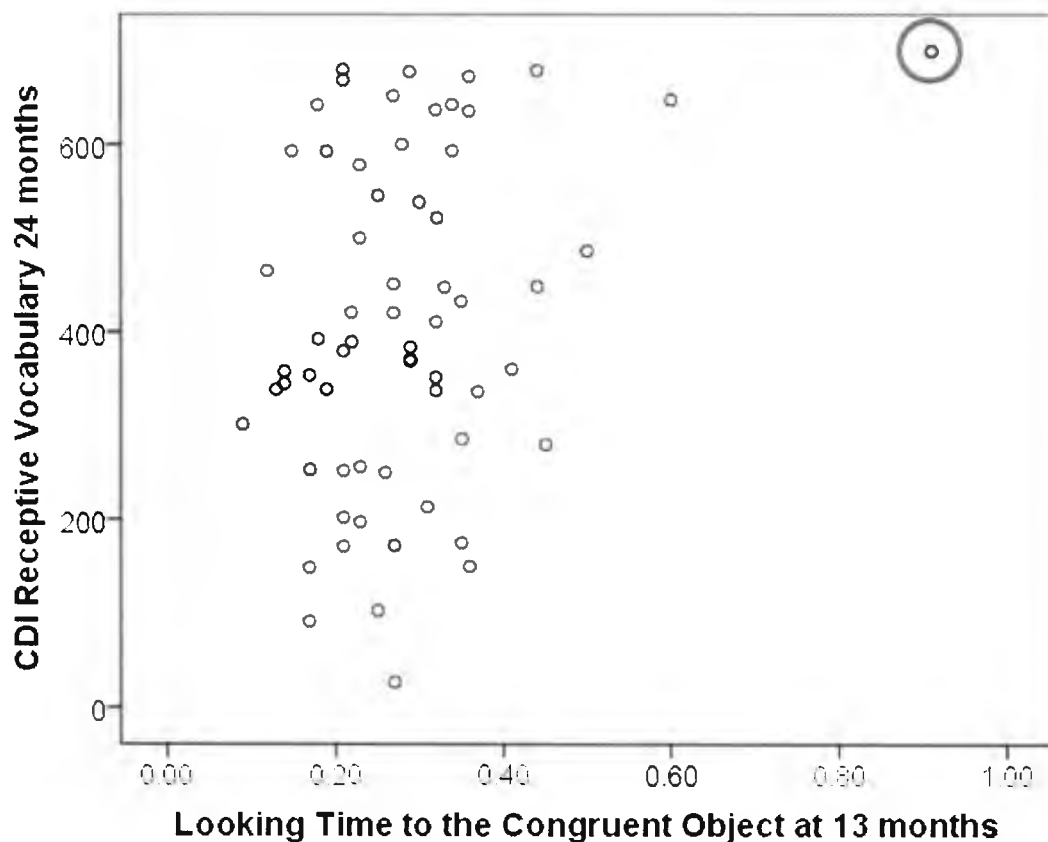


Figure 3.3 Graph showing the relationship between 13 month looking time to the congruent object and CDI receptive vocabulary score at 24 months, with the outlier highlighted.

Running a partial correlation, controlling for 13 month nonverbal IQ score, only marginally reduces the strength of the relationship between looking time and language, $r = 0.25$, $p = 0.047$. Thus it seems that non-verbal IQ has little effect on this relationship.

To look at the effect of group, a multiple regression using the enter method was run, with group and 13 month looking time predicting 24 month receptive vocabulary. Group was a significant predictor ($\beta = -0.306$, $p = 0.01$) with higher receptive vocabulary in the low-risk controls, and looking time became marginally significant ($\beta = 0.21$, $p = 0.09$). Separate correlations were then run in the low-risk and high-risk groups to check whether the relationship between looking time and vocabulary held in both groups. However, the correlations did not reach significance in the separate groups (low-risk controls: $r = 0.27$, $p = 0.1$; high-risk: $r = 0.14$, $p = 0.50$), probably due to the reduced power. However, the strength of the correlation in the control group is similar to that in the overall group and it seems that the relationship between looking time and language is being more strongly driven by the low-risk controls.

3.9 Discussion

Like previous studies, evidence was found for a relationship between early gaze following behaviour and vocabulary. In line with the hypotheses, while increased looking time to the congruent object at 13 months (the measure which was related to ASD and atypical outcomes at 3 years) was associated with higher parent reported receptive vocabulary at 24 months, no significant relationship with vocabulary was found for the first look measure. This suggests that it may be early social understanding, rather than the ability to orient in the direction of another person's gaze, which drives the relationship with later language development.

This prediction of 24 month parent reported receptive vocabulary by early looking time fits clearly with the result reported by Brooks and Meltzoff (2005). The authors argue that, in

their correct first look with vocalisation measure, infants are demonstrating an ability to link a person with an object. This ability implies that infants are showing a level of social understanding relating to the intent of another person. In the gaze following study presented here, it is the measure of looking time, thought to reflect referential understanding, which relates to later vocabulary. Thus it may be an understanding of why people look at objects (i.e., the meaning of the social cue) which facilitates word learning. However, it is also possible that looking time is itself an early measure of social-communicative behaviour in prelinguistic infants. In other words, rather than being a prerequisite for language development, looking time may actually be a measure of the same underlying ability.

Understanding the mechanism driving the relationship between JA and language is an important question, and future studies should aim to further tease apart the different explanations. In this study, the fact that looking time remains a significant predictor of later receptive vocabulary after controlling for the effects of non-verbal IQ from the *MSEL*, suggests that the observed relationship between behaviours is not simply due to both measures reflecting general developmental characteristics. The studies presented in Chapter 4 go further in examining developmental processes, specifically the role of gaze following behaviour in children's ability to learn words.

While there is a significant correlation between looking time and later receptive vocabulary in the overall group, this seems to be driven primarily by the low-risk controls, although neither of the separate group correlations reaches significance. This lack of significance when the groups are split is likely due to the reduced power from a smaller sample size. However, the size of the effect was still much larger in the low-risk control group than in the high-risk group. Thus, while the findings are in line with the literature linking JA to language in typical development (e.g., Baldwin, 1993; Carpenter et al., 1998; Morales et al., 1998; Tomasello & Farrar, 1986), the predicted relationship was not found for

the high-risk infants, although the effect was in the same direction. This may be due in part to reduced variability in the looking time measure among the high-risk infants.

However, it is also important to consider the nature of the language measure used. Both Presmanes et al. (2007) and Sullivan et al. (2007) both used the *MSEL* when linking JA with language in high-risk infants. The *MSEL* measures of language are somewhat different to simple vocabulary measures, and include aspects of behavioural compliance - for example, the RL scale requires the child to point on demand, and follow requests of the experimenter, and the EL language scales requires the child to answer questions. The current study uses the CDI, a parent reported vocabulary measure, which focuses specifically on the total number of words known and spoken by the child, and is thus in some ways a simpler, more intuitive measure than the *MSEL*. Theoretically, orienting in response to gaze might relate to language because it provides a cue to locate the referent of a heard word, and social understanding of the cue is likely to be important for attending to the word-object mapping. Thus, a measure of number of known words, rather than a more complex general developmental measure such as the *MSEL* seems appropriate. However, as this is the first study to relate this particular parent report measure of vocabulary to JA in a high-risk group, there is a need for replication of the results.

Although there was a significant longitudinal association between looking time and later 24 month receptive vocabulary, no evidence was found for concurrent relationships between 13 month looking time and 13 month receptive or expressive language score on the CDI. It is possible that a period of time is required after learning the meaning of eye-gaze cues in order to see the knock-on effects in vocabulary. This lends support to the idea of looking time in the gaze following task being a precursor to language, rather than a measure of the same underlying social-communication behaviour. However, the null effect may also be due to the

fact that language, even comprehension, at 13 months is still very limited and thus perhaps there was not enough variability in our sample.

The relationship between looking time and vocabulary also only held for language comprehension, no relationship was found with a measure of words understood and said, i.e., expressive vocabulary. This is in line with the findings of various other studies and could be due to the fact that language expression is a more complex ability, in that it involves not only an understanding of the word but also a desire to communicate it verbally and motor ability in terms of verbal articulation.

Finally, for first look, no relationship with CDI vocabulary was found at either time point. This suggests that merely following the direction of gaze, at least in this study, is not sufficient for vocabulary development. In the majority of studies using response to JA as the independent variable, no distinction is made between direction of gaze and amount of looking. However, the results here suggest that following another person's eye-gaze does not predict language outcomes per se, and it is the more social-relevant understanding of the meaning of eye-gaze which is important. This has important implications when thinking about the precise behavioural target of an intervention strategy, which will be discussed more fully in the final chapter (Chapter 6).

3.10 Conclusion

In conclusion, gaze following behaviour was decomposed into first look and looking time measures, which are thought to reflect different underlying processes. No risk group differences were found for either measure. However, 13 month looking time to the congruent object, a measure of referential understanding, was significantly reduced in infants who later developed social and communication symptoms of ASD. Further, this looking time measure related to later 24 month receptive vocabulary, an effect driven primarily by the low-risk

infants. This suggests that, as well as difficulties in referential understanding predicting socio-communication problems, understanding of social cues early in development also relates positively to increased vocabulary in typical development.

Summary of Chapter 3

- Gaze following behaviour can be meaningfully decomposed into different measures: first look and looking time.
- No risk-group or clinical outcome group differences were found for the proportion of first looks to the congruent versus incongruent object at either 7 or 13 months of age.
- Looking time to the congruent object was not significantly different between groups at 7 months.
- Looking time to the congruent object was significantly reduced by 13 months in high-risk infants who later develop ASD or other socio-communicative difficulties.
- Reduced looking time to the congruent object at 13-months related to higher ASD symptomatology, as measured by the ADOS-G, at 24 and 36 months of age in the high-risk group.
- Looking time to the congruent object at 13 months was the only measure to show a significant longitudinal relationship with parent reported receptive vocabulary at 24 months.

Chapter 4

Problems in Understanding Social Cues Over Development Underlie Word Learning Difficulties in Children at High Risk for Autism Spectrum Disorder

4.1 Introduction

In the previous chapter, a measure of looking time to the congruent object, thought to index referential understanding, predicted both symptoms of ASD and later vocabulary development. In the current chapter, I examine children's use of social feedback for learning. This goes beyond looking for relationships between looking behaviour and later vocabulary development, and allows the child's use of social feedback to learn a new word to be tested.

The first part of this chapter will present a novel word learning study in 24-month-olds, looking at the use of mutual exclusivity (ME; children's mapping of a novel word to a nameless object, over an object for which they already have a name) and ostensive cues (social feedback e.g., 'yes, this is the *moxi*') to learn new word-object mappings. The relationship between word learning and concurrent 24 month vocabulary on the CDI will also be examined.

The second part of the chapter will look at the relationship between looking time and word learning given social feedback. Both measures are thought to require some degree of social understanding, or 'metarepresentation' to use theory of mind (ToM) terminology. Thus we might expect these measures to be related, with 13 month looking time predicting 24 month word learning. If this is the case, it is possible that learning words using feedback is mediating the observed relationship between gaze following and vocabulary (see Chapter 3).

Part 1: The use of Mutual Exclusivity and Social Feedback to Learn Words in 24-month-olds at High Risk for Autism Spectrum Disorder

4.2 The mutual exclusivity task

The next subsection of this chapter introduces a novel word learning study in 24-month-olds, the ‘mutual exclusivity task’. The task looks at children’s use of mutual exclusivity (ME) and ostensive cues (social feedback) to learn new word-object mappings. The relationship between these measures and concurrent 24 month vocabulary on the CDI is also examined.

4.2.1 Word learning in typical development

When learning new words in a cluttered, dynamic environment, children are faced with the well-known Quinean puzzle– which of the many objects in their visual field is the intended referent of a heard word (i.e., referent indeterminacy; Quine, 1960)? Throughout development, young language learners have various potential strategies that can help them to solve this correspondence problem. Early on in development, when children acquire their first words, ostensive and referential cues are used; the caregiver or experimenter must ostensibly direct the child’s attention to an object and label it repeatedly (Hollich et al., 2000; Houston-Price, Plunkett & Duffy, 2006; Woodward & Hoyne, 1999). Later in development, children can learn words when such cues are not present, by making use of various heuristics to infer a speaker’s referent. One such heuristic, the mutual exclusivity (ME) principle, refers to the assumption that novel words refer to unfamiliar objects or to objects for which the child does not yet have a label. Most children begin using this principle towards the end of the second year of life (Halberda, 2003; Markman & Wachtel, 1988). It is thought that children learn the ME principle by noticing that objects tend to be referred to using only one name (Markman,

1991), and when exceptions occur, such as in bilingual environments, children are less likely to treat object names as mutually exclusive (e.g., Davidson, Jergovic, Imami & Theodos, 1997). Because, in some cases, children are able to make a word-object association after a single labelling episode (Carey & Bartlett, 1978), word learning in this context is referred to as *fast mapping* and has frequently been described as a driving force behind the ‘vocabulary explosion’ which occurs at the end of the second year of life (e.g., Markman, Wasow, & Hansen, 2003).

However, more recently several studies have challenged the central role given to referent selection using ME in children’s vocabulary growth (Horst & Samuelson, 2008; Mather & Plunkett, 2011). In a typical ME task, the child is presented with two or more objects, one of which is an unfamiliar (novel) object, and asked to give or to look at the *moxi* (or another pseudo-word) (Halberda, 2003; Merriman & Schuster, 1991). Correct referent selection (either choosing the novel object or looking longer at it) is in this case taken to reflect correct word learning. However, this differs from the measure of word learning and retention in ostensive word learning situations. In the latter case, looking at the referred object while it is labelled may only reflect cue following and is therefore not considered sufficient evidence of word learning. Correct word-object mapping is typically tested in a separate trial, following the labelling episode, with the child asked to choose the correct referent of a newly learned word amongst two previously labelled objects (Gliga et al., 2012; Houston-Price et al., 2006). When a similar procedure has been used to test word retention following fast mapping, Horst and Samuelson (2008) showed that toddlers who successfully used ME to choose the correct referent of a new word, performed at chance when asked to retrieve that object 5 minutes later. Interestingly, the children’s performance in the retention trials improved if, after their initial correct choice, the experimenter reinforced their knowledge by ostensively labelling the object (i.e., by holding the object while pointing to it and naming it). This suggests that

applying the ME principle may be necessary for quickly finding the referent of a new word but is not sufficient for that word to enter the child's vocabulary. It is feedback on the child's initial choice that seems to be crucial in creating a long-term word-object mapping.

4.2.2 Word learning in ASD

As well as adding to our understanding of word learning mechanisms, these findings have the potential to clarify an interesting paradox in language acquisition in ASD. Children with ASD are less responsive to social cues, particularly referential cues and experimental studies have shown that they have difficulties using these cues for word learning (Baron-Cohen, Baldwin, & Crowson, 1997). However, referent selection using ME seems to be intact in ASD (de Marchena, Eigsti, Worek, Ono & Snedeker, 2011; Preissler & Carey, 2005). This seems surprising given that children with ASD have smaller vocabularies than expected for their age (Charman, Drew, Baird & Baird, 2003; Hudry et al., 2010; Tager-Flusberg, Paul & Lord, 2005) and that delays in language acquisition (not accompanied by non-verbal gestural communication) form part of the diagnostic criteria for ASD (ICD-10; World Health Organisation, 1993). There are two possible explanations for this discrepancy. Firstly, it could be that studies have overestimated these children's referent selection abilities. Because ASD is rarely diagnosed before 2 years of age, ME has mostly been assessed in older children (Preissler & Carey, 2005 tested 5- to 9-year-olds; de Marchena et al., 2011 assessed children aged 7-11 years). Their ability to use the ME constraint might have been the result of an extended learning process and therefore not a contributor to word learning earlier in development. Despite their age and proven ME skills, average comprehension vocabulary in Preissler and Carey's (2005) sample was only equivalent to that of a typical 2-year-old. A second explanation for the discrepancy between vocabulary size and word learning strategies may be explained by possible word retention difficulties. It could be that children with ASD,

who are less sensitive to social-communicative cues, do not benefit from ostensive feedback and are therefore less able to retain the word-object mapping, despite their demonstrated ability to use the ME bias.

4.2.3 Word learning in infants at high risk for ASD

To tease apart these hypotheses, we replicated Horst and Samuelson's (2008) study with a sample of 24-month-olds who were either at high- or low-risk for ASD. The few studies that have examined language development in high-risk children show that they are slower to acquire language (Toth, Dawson, Meltzoff, Greenson & Fein, 2007; Yirmiya et al., 2006; Yirmiya, Gamliel, Shaked & Sigman, 2007). There are, however, even fewer studies that have investigated word-learning strategies in this population. We know that high-risk 3-year-olds have difficulties using ostensive and referential cues to learn a new word-object mapping (Gliga et al., 2012), but little is known about available strategies earlier in development. Studying 24-month-olds at high risk for ASD enables us to investigate whether the linguistic difficulties measured in this population are related 1) to difficulties with referent selection through ME, apparent earlier in development (but not later in life; Preissler & Carey, 2005); or 2) to difficulties using feedback for long-term word retention. First, children's ability to fast map a novel word to an unfamiliar object was assessed. Following each trial choice, we either provided no feedback, or ostensively labelled the novel object, thus correcting or reinforcing the child's initial choice. Memory for the word was retested after a 5 minute break.

4.3 Method

4.3.1 Participants

Thirty-one toddlers at high risk for ASD and 44 low-risk children took part in this study. Six additional children (1 low-risk, 5 high-risk) participated but were not included in the analysis due to non-compliance. To take account of the non-verbal IQ differences ($t(71) = 3.28$, $p = 0.002$), data for 13 low-risk toddlers were removed, including two children who had no *Mullen Scales of Early Learning (MSEL)* (Mullen, 1995) data and the 11 children with the highest non-verbal scores. Following exclusion, *MSEL* non-verbal scores did not differ significantly between the groups ($t(52.5) = 1.68$, $p = 0.1$). Participants included in the final analysis were 31 low-risk toddlers (13 boys and 18 girls, mean age = 24.3 months, $SD = 0.59$) and 31 high risk toddlers (14 boys and 17 girls, mean age = 24.6 months, $SD = 1.02$).

4.3.2 Stimuli

Stimuli for the word-learning task were 16 familiar objects, eight of which were designated 'target familiar' (spoon, toy duck, key, toy horse, ball, toy car, baby shoe, toy pig) and eight of which were designated 'non-target familiar' (toy cow, cup, toothbrush, pen, hairbrush, book, fork, comb). An additional eight novel objects were similarly either 'target novel' or 'non-target novel' (egg poacher, bottle stopper, lemon juicer, avocado slicer, bottle opener, cooking brush, fried egg shaper, whisk). The 'target familiar' objects were chosen based on the normative MacArthur Bates Communicative Development Inventory (CDI) estimates for 24-month-olds (Dale & Fenson, 1996) with all object labels reported to be known by at least 76% of toddlers at that age. For the 'target novel' objects, four novel, bi-syllabic, pseudo-words were used: *moxi*, *fimit*, *kela* and *togo*. The 24 objects available were then split into groups of three for each trial, with two familiar objects, and one novel object. Each child completed eight trials, four with 'target familiar' items (Familiar trials) and four

with a ‘target novel’ item (Novel trials). In the Familiar trials, a ‘target familiar’ object was paired with another ‘non-target familiar’ object and a ‘non-target novel’ object. In the Novel trials, the ‘target novel’ object was paired with two ‘non-target familiar’ items. Care was taken to ensure that none of the ‘non-target familiar’ objects names were phonologically close to the name used for the novel object. Two alternative options for the assignment of groups of objects to the trial types was created and use of one or the other option was counterbalanced across children in both groups. The objects were presented to children on an unpartitioned rectangular tray.

4.3.3 Procedure and design

The study was split into three phases: *familiarisation* with the objects, *fast-mapping*, and, following a 5 minute delay, *retention* (see Figure 4.1). During the initial familiarisation period, children played with all of the objects for 5 minutes to ensure that novelty preference would not interfere with children’s later choices in the experimental task (Mather & Plunkett, 2010). An experimenter made sure that the child saw all of the objects by asking the child to place the objects one by one into a box. No objects were named, during this phase, by either the experimenter or the parent. The experimental task then began, with the child seated at a small table, either alone or on the parent’s lap, facing the experimenter.

8 x Fast-mapping

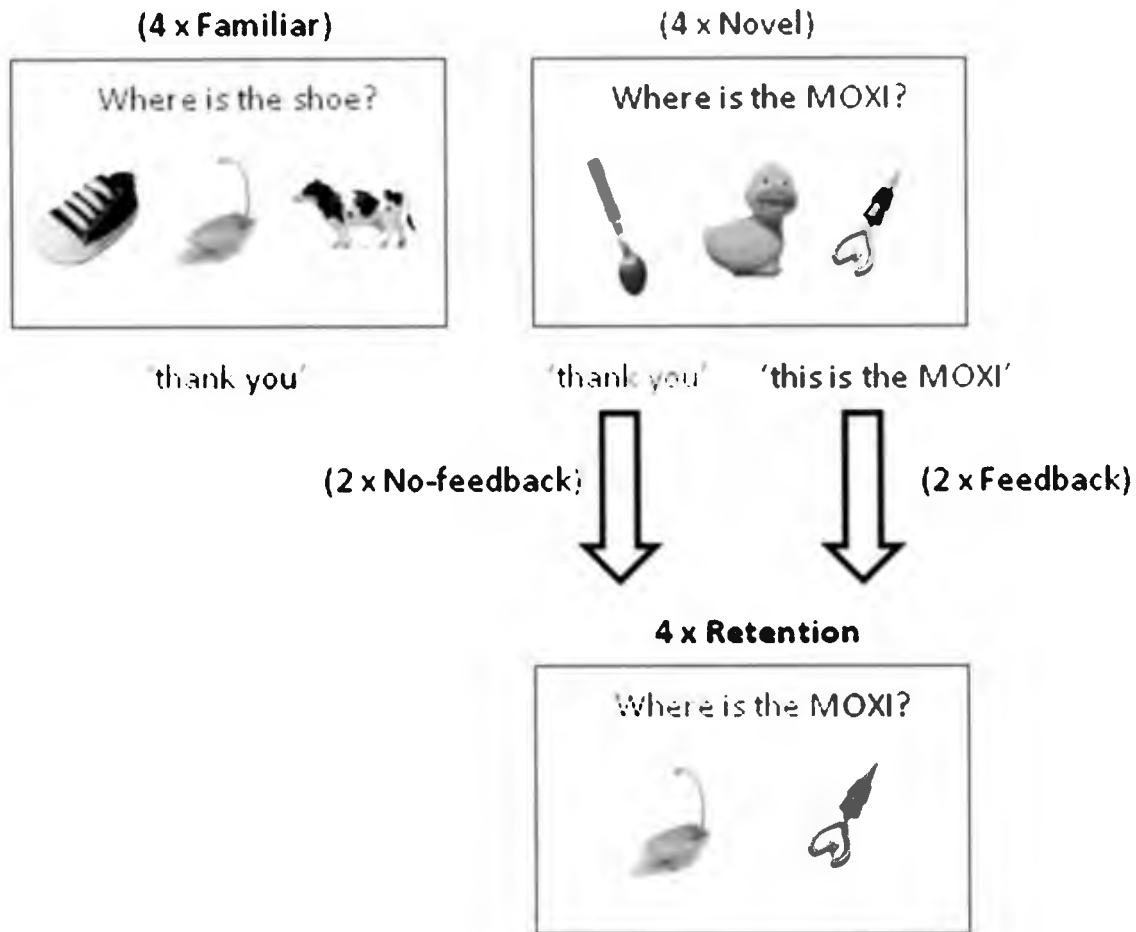


Figure 4.1 Experimental paradigm. Example of objects and words used in fast-mapping and retention trials. * indicates performance significantly above chance at the $p < 0.05$ level.

Fast-mapping trials: At the start of each trial the experimenter held the tray of objects and, looking at the child, said 'Can you see the spoon/moxi?' The tray was then placed in front of the child and the experimenter asked 'Can you give me the spoon/moxi?' On all Familiar trials and on two of the Novel trials the experimenter responded 'Thank you', irrespective of which object the child chose. These were therefore 'No-feedback' trials. In the remaining two Novel trials (i.e., Feedback trials) the experimenter ostensibly labelled the correct object, holding it in front of the child and responding either 'Yes, this is the moxi. What a nice

moxi!’ or ‘No, this is the moxi. What a nice moxi’ depending on whether the child’s choice had been correct or incorrect.

Retention trials: Following a 5 minute break during which children played in the testing room with other toys, there were four retention trials. Pairs of only Novel objects were presented, each including a ‘target novel’ object and a ‘non-target novel’ object (which had been seen during familiarisation and the Familiar fast-mapping trials, but which had never been named). Previously target novel objects were paired with non-targets so that memory for all four target objects can be tested independently. For two of the retention trials, the ‘target novel’ object had previously been ostensively labelled during fast-mapping (Feedback trials), while the other two trials had included no labelling (No-feedback trials). Retention trials again followed one of two pre-determined orders, with selection counterbalanced across children.

4.3.4 Data analysis

Children’s responses were video coded during both the fast-mapping and retention trials. If the child made no response (i.e., did not touch or give any of the objects), then the trial was discarded as invalid. Valid trials were coded as correct or incorrect on the basis of the object given by the child to the examiner (or, if no object was given, then on the basis of the first object touched by the child). A second coder rated trials for 4 high-risk (13%) and 4 low-risk (13%) toddlers, yielding 100% agreement between coders.

Data were analysed with mixed models ANOVA. This data violates the normality assumptions of ANOVA. However, I was interested specifically in the between subjects-within subjects interaction effect, for which there is no non-parametric equivalent. Given that under non-normal data ANOVAs are *less* likely to give a significant effect (increase in type II error) any significant findings should be robust.

4.3.5 Measures of language and general development

General developmental level was assessed at the same visit, using the *MSEL* (low-risk $n = 31$, high-risk $n = 31$). A non-verbal composite score was calculated as the average of the Visual Reception and Fine Motor scale T-scores. The CDI (Fenson et al., 1993) a parent report measure of vocabulary, was also collected for toddlers in both groups (low-risk $n = 30$, high-risk $n = 26$). A receptive vocabulary count was calculated by combining the total numbers of words ‘understood’ and words ‘understood and said’ for each child. Group characteristics are shown in Table 4.1.

Table 4.1 Descriptive Statistics for the Mullen Scales of Early Learning (*MSEL*) and MacArthur Bates Communicative Development Inventory (CDI) Receptive Vocabulary

	Low-risk ($n = 31$)		High-risk ($n = 31$)	
	M	SD	M	SD
Age	24.3	0.59	24.6	1.0
Female:Male	18:13		17:14	
CDI	(n = 30)		(n = 26)	
Receptive vocabulary count	449.0	172.1	335.2	166.8*
<i>MSEL</i>	(n = 31)		(n = 31)	
Non-verbal T-score	53.58	7.03	49.79	10.45
Verbal T-score	57.55 (7.74)		50.12 (12.39)*	

4.4 Results

The fast-mapping and retention trials are analysed separately. In each case, a mixed-factorial ANOVA is used to test for differences in word learning performance as a result of Group (varying between-subjects; low-risk, high-risk), and fast-mapping Item type (varying within-subjects; Novel, Familiar) or retention Feedback type (also varying within-subjects; Feedback, No-feedback). The number of valid trials did not differ significantly between Groups (Novel item: High-Risk $M = 3.9$, Low-Risk $M = 4.0$; Familiar item: High-Risk $M = 3.9$, Low-Risk $M = 4.0$; Ostensive feedback: High-Risk $M = 1.9$, Low-Risk $M = 1.9$; No-feedback: High-Risk $M = 2.0$, Low-Risk $M = 1.9$). Performance is also compared to chance levels to highlight for which group and under which conditions participants successfully choose the correct referent or remembered its label. The retention data were subsequently re-analysed by separating those trials for which ostensive feedback provided either a correction of an initially incorrect choice or reinforcement of an initially correct choice.

Because of the greater number of girls in the low-risk group and the frequently reported superiority of girls' language skills (observed also in this study with girls' average receptive vocabularies of 471 words, whereas for boys this was 295 words; $t(54) = -4.2$, $p < 0.001$), I initially entered Gender as a between-subjects factor in all statistical analyses. However, Gender did not affect the significance level of any main factors of interest (Trial, Group), nor did it significantly interact with them. As such, I have collapsed across gender in the analysis presented here.

4.4.1 Fast-mapping

One sample t-tests against a chance level of .33 showed that both low-risk and high-risk toddlers performed significantly above chance in selecting the correct object in both Novel and Familiar fast-mapping trials (low-risk novel: $t(30) = 3.41$, $p = 0.002$, low-risk familiar:

$t(30) = 25.78, p < 0.001$; high-risk novel: $t(30) = 3.88, p = 0.001$, high-risk familiar: $t(30) = 9.77, p < 0.001$, see Figure 4.2).

A mixed-factorial ANOVA revealed a significant main effect of Trial ($F(1,60) = 84.20, p < 0.001, \eta^2 = 0.58$), with both groups showing superior performance on Familiar trials, but no significant main effect of Group ($F(1,60) = 0.17, p > 0.1, \eta^2 = 0.003$). High-risk toddlers performed slightly worse than controls on familiar item trials and slightly better than low-risk controls on novel item trials, but the Trial by Group interaction was marginally significant ($F(1,60) = 3.47, p = 0.07, \eta^2 = 0.06$), with only a small effect size.

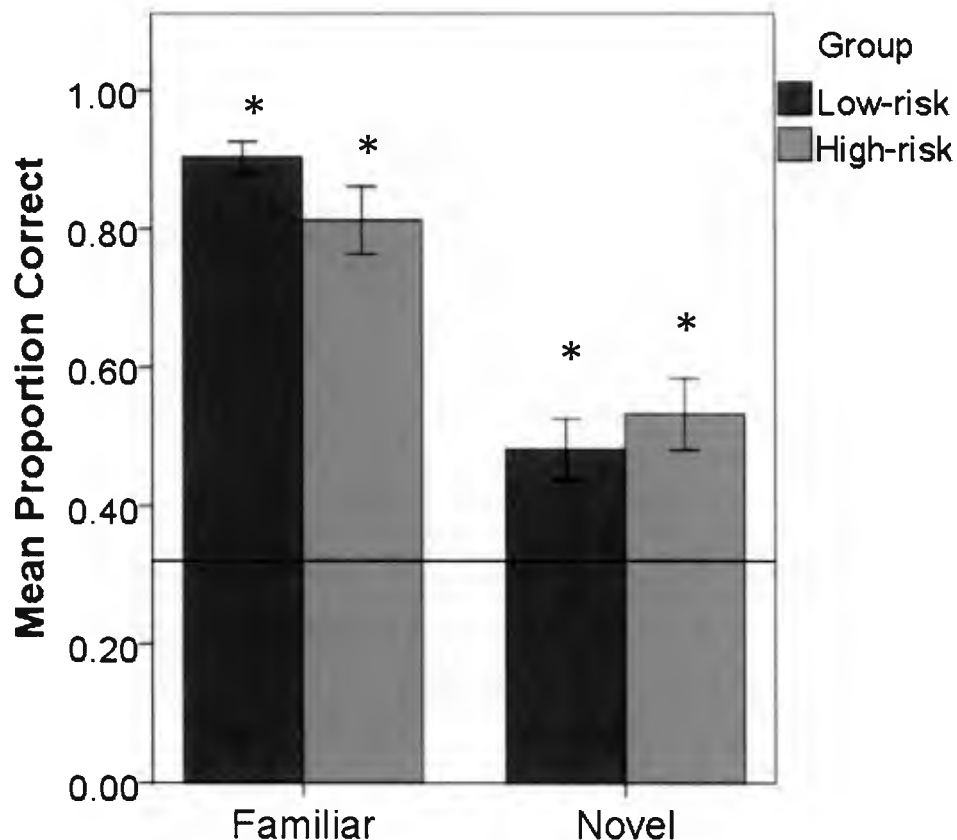


Figure 4.2 Performance during the fast-mapping trials. Chance level is 0.33. Error bars +/- 1 standard error. * indicates performance significantly above chance at the $p < 0.05$ level.

4.4.2 Retention trials

Retention trials were split into those for which children had initially received Feedback or No-feedback during the fast-mapping phase (Figure 4.3). Twenty-eight high-risk toddlers and 28 low-risk toddlers contributed data to this measure. One-sample t-tests against a chance level of .5 revealed that only low-risk children in the Feedback trials performed significantly above chance, ($t(27) = 4.4$, $p < 0.001$).

A mixed-factorial ANOVA with Feedback Type (Feedback vs. No-feedback) and Group yielded a significant main effect of Group, ($F(1,54) = 5.68$, $p = 0.02$, $\eta^2 = 0.1$), with low-risk children performing better overall than the high-risk group. The main effect of Trial was also significant, reflecting an increased proportion correct on the trials with Feedback ($F(1,54) = 5.36$, $p = 0.02$, $\eta^2 = 0.1$). There was no significant interaction between Trial and Group ($F(1,54) = 0.86$, $p = 0.36$, $\eta^2 = 0.02$).

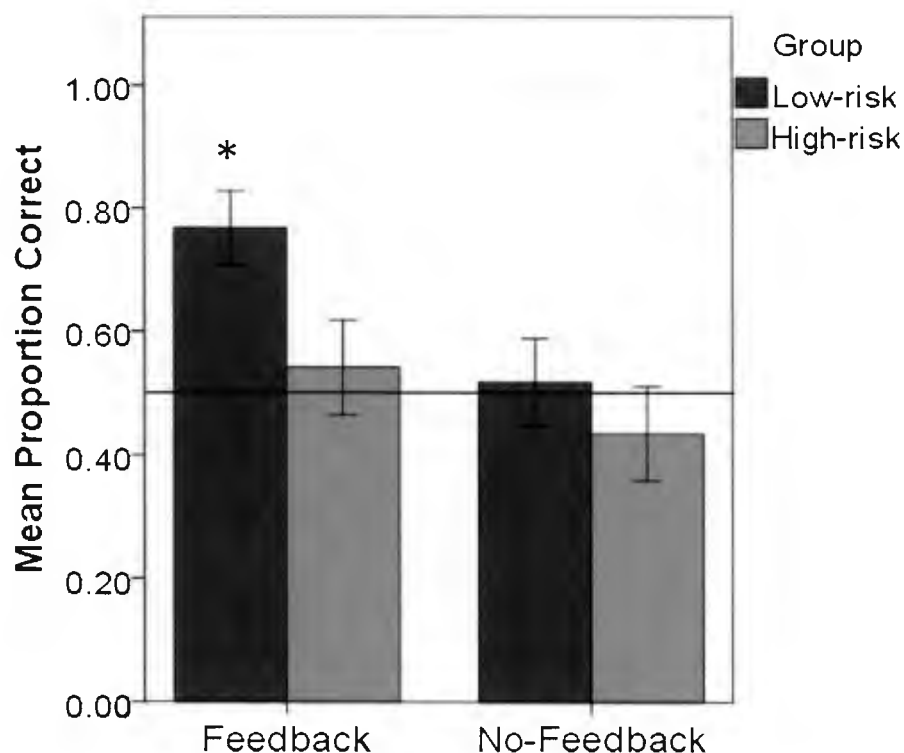


Figure 4.3 Performance during the retention trials. Chance level is 0.50. Error bars +/- 1 standard error. * indicates performance significantly above chance at the $p < 0.05$ level.

4.4.3 Type of feedback

Not all participants contributed both correct and incorrect trials (some participants were always correct or always incorrect in their choices). Because ANOVAs incorporating both Feedback type (Feedback versus No-feedback) and Initial response (correct, incorrect), could not be run, I decided to estimate the effect of Feedback Type on initially correct (i.e., reinforced) and initially incorrect (i.e., corrected) trials, separately. Different numbers of participants contribute data to these analyses.

In order to determine whether memory for a word was above chance performance, one-sample t-tests were carried out, looking at whether Feedback or No-feedback on choices that were initially correct or incorrect resulted in above chance memory for the word. Data from 20 low-risk and 10 high-risk toddlers contributed initially incorrect trial data. Eighteen low-risk and 19 high-risk participants contributed data for the analysis of initially correct choices. Both low-risk and high-risk infants were at chance performance when given no feedback, irrespective of their initial choices. Low-risk children showed above chance retention performance given feedback, for both correction ($t(19) = 3.0, p = 0.008$) and reinforcement ($t(17) = 2.4, p = 0.03$) of their initial choice. High-risk infants performed better when their initial choice was correct with a trend in the direction of above chance performance, although this did not reach significance ($t(18) = 2.0, p = 0.06$). However, the high-risk infants did not perform above chance following correction ($t(9) = -1.9, p = 0.1$).

When directly comparing the effect of Feedback on initially incorrect trials, a significant main effect of Group was found ($F(1,28) = 8.4, p = 0.007, \eta^2 = 0.23$) suggesting that the retention of incorrect fast-mapped trials was problematic for the high-risk toddlers. Although high-risk toddlers performed particularly poorly after Feedback (see Figure 4.4) the interaction between Group and Feedback Type ($F(1,28) = 3.44, p = 0.07, \eta^2 = 0.11$) was

marginally significant, with only a small effect size. The same analysis applied to initially correct trials yielded no main effect (of Group or Feedback type) nor a significant interaction between Group and Feedback type.

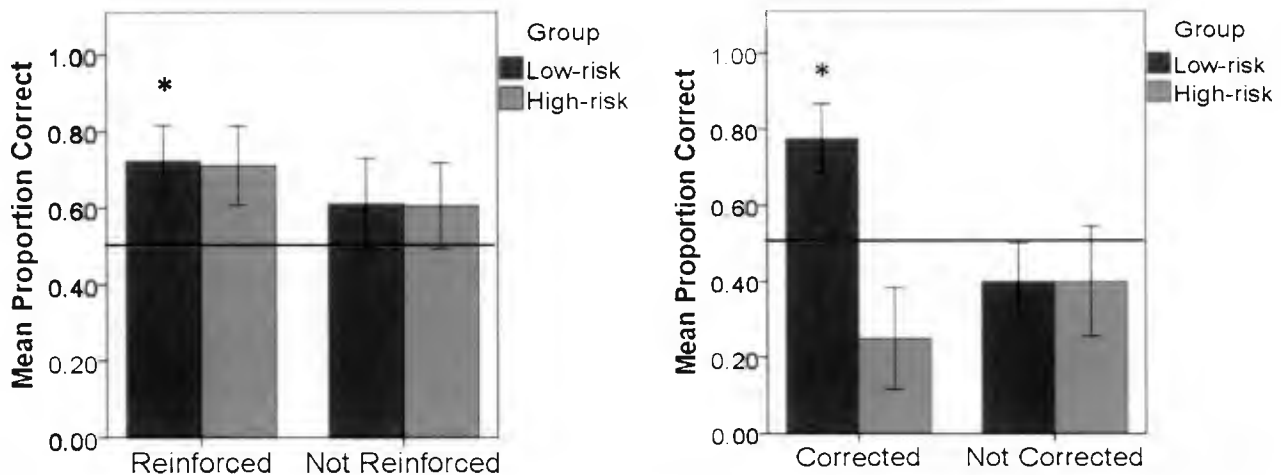


Figure 4.4 Performance during the retention trials, depending on whether the feedback was reinforcing or correcting children's initial choices. Chance level is 0.50. Error bars +/- 1 standard error. * indicates performance significantly above chance at the $p < 0.05$ level.

4.4.4 Relationship with receptive vocabulary size

Finally, I was also interested in exploring which of the abilities measured in the current experimental task might relate to linguistic competency as measured by a standardised assessment of language ability; the CDI. As seen in Table 4.2, there was a correlation overall between performance in the Feedback trials and receptive vocabulary count. When this correlation was broken down by Group, only performance of the high-risk children during the Feedback trials was positively correlated with receptive vocabulary. Our study demonstrates that in high-risk children, the ability to use feedback is associated with receptive vocabulary. The lack of a correlation in the low-risk group is probably due to the smaller variability in task performance.

Table 4.2 Correlations between Experimental Performance and Vocabulary Size

CDI	Familiar	Novel	Feedback	No-feedback
Receptive				
<i>Overall</i>				
<i>Corr. coef</i>	0.08	0.19	0.31*	0.12
<i>p</i>	ns	ns	.03	ns
<i>n</i>	56	56	50	50
<i>Low-risk</i>				
<i>Corr. coef</i>	0.13	-0.11	0.09	0.29
<i>p</i>	ns	ns	ns	ns
<i>n</i>	30	30	27	27
<i>High-risk</i>				
<i>Corr. coef</i>	0.18	0.32	0.44*	-0.13
<i>p</i>	ns	ns	.04	ns
<i>n</i>	26	26	23	23

4.5 Discussion

The high-risk children in this sample showed significantly smaller receptive vocabulary size than the low-risk controls (see Table 4.1). This pattern is the same as that found by Hudry et al. (2010) in a sample of young children with an ASD diagnosis. As previously

demonstrated for older children with ASD (de Marchena et al., 2011; Preissler & Carey, 2005), toddlers at high risk for ASD had no problems, in this study, using ME to find the referent of a new word. Their performance could not be discriminated from that of low-risk controls in either the Novel or Familiar word trials. While correct fast-mapping in both groups was lower than the rates found by Horst and Samuelson (2008) this may be due to the fact that the task was administered mid-way through a battery of other assessments.

Given no feedback on their initial choice, both groups of children performed at chance when asked to retrieve the correct referent of a newly learned word, 5 minutes later, replicating earlier findings by Horst and Samuelson (2008). However, while low-risk toddlers benefited from feedback, which brought their retention performance to a level above chance, high-risk toddlers did not show this effect. This poor performance is even more striking given that the retention task was less demanding than that used by Horst and Samuelson (2008). In their task, three objects were used at test, two of which had been previously labelled. The fact that high-risk toddlers could apply the ME principle suggests that the successful performance of older children with ASD shown by Preissler and Carey (2005) is not the result of a gradual learning process. This strategy seems to be available from early on in vocabulary acquisition. In contrast to their good performance in the fast mapping trials, high-risk toddlers did not benefit from feedback for the long-term retention of word meaning.

Performance during retention trials suggests that high-risk toddlers are generally less able than low-risk controls to remember which object had previously been labelled. A separate analysis of Corrected and Reinforced choices revealed that high-risk children showed difficulties with using *corrective* feedback to update initial incorrect word-object mappings. For initially incorrect trials, although there was no significant interaction between Trial (Feedback versus No-feedback) and group, only low-risk children showed above chance

performance when being corrected. In their study of typically developing toddlers, Horst and Samuelson (2008) only analysed retention of initially correct choices (which form the majority of choices). However, incorrect fast mapping is probably not a real-life rarity in word learning, and feedback becomes of even greater importance in this situation. The literature on how children update initially incorrect choices is still scarce. Interestingly, and against expectations, in a study investigating learning in typically developing older children, performance in a fact-acquisition task was not affected by initial incorrect guessing (Kang et al., 2011), suggesting that self-generated hypotheses are typically easily overwritten by explicitly taught information, at least in typical populations. The same can be seen here in low-risk controls, who reach similar performance levels after feedback, for both initially correct and incorrect choices (Figure 4.4).

It is important to consider what factors could explain high-risk toddlers' lesser reliance on feedback in general and on corrective feedback in particular. The type of intervention children received upon their initial choice confounded feedback with ostensive labelling, any of the two being potentially problematic in ASD. Children and adolescents with ASD are often described as lacking cognitive flexibility (Goldstein, Johnson & Minshew, 2001; Kleinhans, Akshoomoff & Delis, 2005) which could prevent them from easily updating information. However, recent studies using card sorting, where the sorting rule had to be updated repeatedly, did not put children with ASD at a disadvantage (Dichter et al., 2009; Poljac et al. 2010). Alternatively, toddlers at high risk for ASD might give less weight to information gained from others, through labelling, than to self-generated hypotheses. Children's object choices in the ME paradigm are the result of a hypothesis testing strategy, where each hypothesis (i.e., possible novel-word referent) is assigned a certainty level (Horst, Scott & Pollard, 2010). The corrective effect of feedback, in typical populations, may therefore be due to an increased certainty for an ostensively labelled object referent.

Difficulties with social interaction, which manifest as either decreased responsiveness to social cues or a lack of initiation of social interaction, are defining characteristics of ASD (Lord et al., 2000) and children at high risk for ASD have also been shown to rely less on ostensive and referential cues for word learning (Gliga et al., 2012; Parish-Morris, Hennon, Hirsh-Pasek, Golinkoff & Tager-Flusberg, 2007). It is therefore possible that the high-risk toddlers in this study ignored the socially conveyed information, or did not appreciate the certainty value of the ostensive feedback they received regarding their choices. The current study cannot tease apart these two possibilities i.e., whether results are due to a lack of child flexibility in word learning, or differential weighting of self-generated versus taught information. Future studies will have to address this by comparing the effects of ostensive versus non-ostensive corrective evidence.

These findings are timely because they shed light on the role played by referent selection strategies on vocabulary learning in general. Lexical constraints or heuristics like ME have been described as the cause of the acceleration in word acquisition taking place towards the end of the second year of life. Markman, one of the first researchers to explore children's use of ME, states that: 'At some point the learning changes and becomes very rapid. This new fast form of learning may be made possible by the emergence, consolidation or learning of such constraints on word learning' (Markman, 1991, p. 80). However, recent computational models of vocabulary growth do not support *the need for* specialised processes (McMurray, 2007). The findings presented here also suggest that the ability to apply the ME principle is not *sufficient* for vocabulary growth. While hearing a new word in the presence of only one novel object might serve to move word learning forward, through processes of hypothesis-generation regarding possible referents, these hypotheses must either be re-tested or otherwise directly confirmed through feedback in order to result in successful word learning.

It is retention performance following feedback rather than fast-mapping performance that correlates with children's receptive vocabulary size.

Although causation cannot be inferred from correlations, it is unlikely that lower retention in the high-risk group is due to lower language scores. All participants were matched on non-verbal IQ because memory itself (ability tested by the visual reception scale of the *MSEL*) could have affected performance in the retention test. All children, however, could follow simple commands like 'Can you give me the moxi?' and further, both high and low-risk children showed excellent performance in the fast-mapping trials. I therefore believe it is more likely that children's ability to use feedback is driving their vocabulary acquisition, explaining partially the difference in vocabulary size between low and high-risk participants. Although this study shows that the ability to learn from feedback contributes to word retention, it is unlikely that this factor is unique in explaining vocabulary growth in children with ASD or in those at high risk for the disorder. The child's willingness to take part in social interaction, their request for lexical information, and their ability to extract lexical rules, all of which are problematic in these populations (e.g., Eigsti, de Marchena, Schuh & Kelley, 2011; Tek, Jaffery, Fein & Naigles, 2008) are probably contributing factors. Future studies will have to assess their relative contribution for language acquisition in typical and atypical development.

*Part 2: Developmental Continuity in Problems Understanding Social Cues in Children at
High Risk for Autism Spectrum Disorder*

4.6 Gaze following and mutual exclusivity

Both looking time to the congruent object in the gaze following task and word learning using feedback in the mutual exclusivity task are thought to reflect some level of social understanding. Further, these are the measures that relate to ASD outcome/risk status, and 24 month receptive vocabulary (Chapters 3 & 4). Here, I aimed to test whether looking time, which is an index of referential understanding, is a precursor to later learning from social feedback. If looking time does relate to later use of social feedback to learn words, then this could offer a potential mediating mechanism for the observed relationship between gaze following and vocabulary.

According to the Theory of Mind (ToM) account (Baron-Cohen, Leslie & Frith, 1985), discussed in Chapter 1, typically developing children possess metarepresentational understanding, which is manifested early in development in joint attention (JA), and later in the passing of false belief tasks. The ToM account argues that metarepresentation, or understanding that other people have their own minds, is impaired in children with ASD. Baron-Cohen (1989a) makes the distinction between following a person's gaze and understanding that the person is interested in the object that they are attending to, with only the latter requiring a metarepresentational understanding. This distinction is reflected in the measures from the Chapter 3 gaze following task, with reduced looking time to the congruent object in the infants who later show social-communication difficulties compared to low-risk controls.

If there is indeed a continuum of skills involving metarepresentation, it would be predicted that looking time, which is thought to index referential understanding, should relate to later use of social information for learning words. In the mutual exclusivity task, feedback trials involve the experimenter holding the correct object in front of the child and saying ‘this is the moxi’. This type of feedback is ‘social’ in the sense that it comes from another person, although the precise mechanism by which this feedback contributes to the word-object mapping is not clear. Several factors are likely to play a role, including repetition of the object name and social cues increasing the salience of the object (Samuelson and Smith, 1998). Further, as discussed earlier in this chapter, feedback trials can be decomposed into those that were correct initially (reinforced) or incorrect initially (corrected). The mechanisms for learning from reinforcing feedback versus correcting feedback may operate somewhat differently. While learning from both reinforcement and correction involves a degree of social understanding, correction trials also require flexibility in switching response.

Whether a relationship between looking time and fast mapping performance would be predicted depends on how ME is operating. If, as proposed by social-pragmatic theory (Tomasello, 2003), children infer the speaker’s intent in order to disambiguate the intended referent, then fast mapping involves social understanding and should be related to earlier referential understanding in the gaze following task. However, if ME is a lexical heuristic, i.e., each object is referred to by a single name, as implied by the fact that bilingual children do not show ME (e.g., Davidson et al., 1997), then no relationship with earlier looking time would be predicted.

It was hypothesised that looking time at 13 months would predict 24 month word learning using feedback, and that this relationship would be primarily driven by performance on reinforced trials. I also aimed to test whether there was a relationship between looking time

and fast-mapping. A significant relationship would be consistent with the social-pragmatic account rather than ME as a lexical heuristic.

4.7 Method

Data analysis

As discussed above, the word learning data are not normally distributed, so the word-learning measures were converted into categorical variables (see Table 4.X).

Fast-mapping trials: There were a range of different responses for fast mapping trials, 0, 0.25, 0.33, 0.5, 0.67, 0.75 and 1. I therefore created an ordinal variable with 0 (< 0.5), 1 (for 0.5), 2 (for > 0.5).

Feedback trials: Another three category ordinal variable was created (0, 1 and 2) following the same criteria as for fast-mapping trials.

Reinforced and corrected trials: There were less data for the reinforced and corrected trials so to avoid having a tiny number of children in each category, responses were split to create binary variables, with '1' for all correct, and '0' for not all correct.

Table 4.3 Frequencies for categorical word learning variables

Categorical score	Fast-mapping N = 75	Feedback N = 56	Reinforced N = 37	Corrected N = 30
0	19	10	13	13
1	26	19	24	17
2	30	27	-	-

4.8 Results

First, a saturated path analysis model using a mean and variance adjusted weighted least squares (WLSMV) estimator was run with looking time at 13 months predicting 24 month word learning using feedback and fast mapping (see Figure 4.5, shows unstandardized estimates). All three variables (looking time, feedback and fast mapping) were regressed on group. The model showed no significant relationship between looking time and either fast mapping performance (odds ratio; OR = 0.22, confidence interval; CI = 0.03 – 1.45, $p = 0.12$) or word learning from feedback (OR = 1.03, CI = 1.30 – 35.70, $p = 0.43$). This lack of a significant relationship between looking time and word learning using feedback means that learning from feedback is not a candidate for a mediator variable in the relationship between looking time and 24 month vocabulary (Chapter 3).

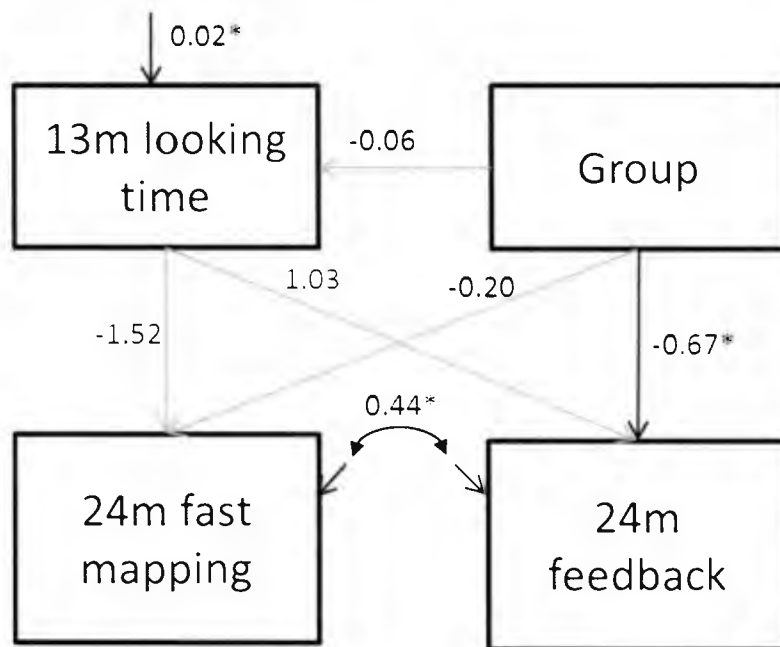


Figure 4.5 Path analysis model looking at the relationship between 13 month looking time and 24 month fast mapping and feedback performance, controlling for risk status.

Given the hypothesis that the relationship between looking time and word learning would be primarily driven by performance given reinforcement, a similar model, but with the binary

reinforcement and correction variables instead of feedback was run (see Figure 4.6). There was a significant relationship between group and performance on correction trials ($OR = 0.25$, $CI = 0.08 - 0.73$, $p = 0.01$), with higher score on correction trials more likely in the low-risk control group, as expected from the results in the first part of this chapter.

As in the previous model (Figure 4.5), there was no significant relationship between looking time and 24 month fast mapping ($OR = 0.22$, $CI = 0.03 - 1.46$, $p = 0.12$). However, in line with the hypotheses, 13 month looking time significantly predicted word learning from reinforcement ($OR = 109.73$, $CI = 9.96 - 1208.38$, $p < 0.001$), with longer looking time predicting increased odds of choosing the correct object after reinforcement, but not correction ($OR = 7.53$, $CI = 0.11 - 535.91$, $p = 0.35$).

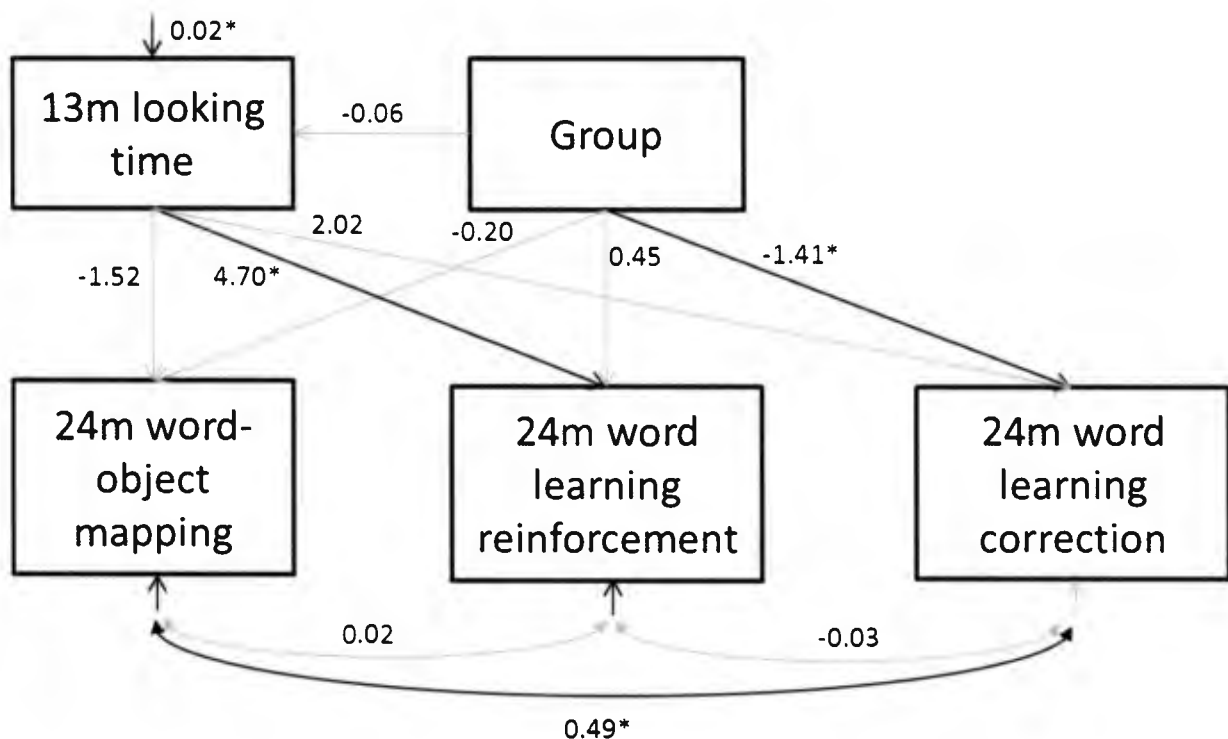


Figure 4.6 Path diagram of the relationship between 13 month looking time behaviour and mutual exclusivity and later word learning following reinforcement and correction.

4.9 Discussion

The results provide some support for the predicted link between early gaze following behaviour and later use of social feedback to learn words. While, contrary to the hypothesis, a significant relationship between 13 month looking time and learning words from feedback was not found, looking time did predict later learning from reinforcement (when the child's initial object choice was correct, and feedback reinforced their selection). Further, no relationship with the correction trials, which were initially incorrect, was found. Finally, in line with the idea that mutual exclusivity is a lexical heuristic, the models did not show a significant relationship between looking time and later fast-mapping performance.

The lack of a significant relationship between fast-mapping ability and earlier looking time, suggests that use of ME may indeed reflect a more 'low-level' lexical principle. This finding is in contrast to social-pragmatic theory, which argues that ME involves inference of the speaker's intent (i.e., after being asked for the 'moxi', the child reasons if he/she wanted the duck, rather than the novel object he/she would have said 'duck' because that is what we both call that object, therefore he/she must mean the new object). However, as discussed earlier, Markman (1991) argues that ME actually develops due to children learning that objects tend to be referred to by a single name. This is more akin to statistical learning or contingency learning, where infants must extract patterns of information. Even very young infants, in the first few months of life, are able to extract statistical regularities, for example when learning how continuous speech is segmented into different words (Saffran, Aslin & Newport, 2008). This fits with MacWhinney's (1989) claim that ME reflects a useful learnt heuristic, rather than an innate constraint on word learning.

The lack of a significant relationship with word learning from feedback may be due to the fact that, for trials which were initially incorrect, the child is required to both understand the

importance of the social information and use it to re-map the word to a different object. Perhaps, the requirement for cognitive flexibility is more important to success in these trials, rather than social understanding per se. This idea is supported by the fact that no relationship between looking time and correction trials was found, but there was a significant relationship for performance on reinforcement trials.

When the child's choice was initially correct, reinforcement offers a social confirmation. The social cue from another person indicates their interest in the object, potentially enhancing the salience or importance of the word-object mapping for the child. The relationship between 13 month looking time and learning from reinforcement is in line with the predictions from the theory of mind account. Early difficulties in understanding that other people have their own beliefs and desires is likely to have knock on consequences in a variety of tasks, involving such social understanding, across development.

It seems, from the results presented in Chapters 3 and 4, that there are early behavioural differences in understanding socially-relevant information in high-risk infants and those who later develop symptoms of ASD. An EEG study (Elsabbagh et al., 2012), with the same group of infants aged 6-10 months, has shown differences between the low-risk controls and high-risk infants who later develop ASD, in terms of their neural response to eye-gaze cues. Infants who go on to an ASD diagnosis at 3 years show less discrimination of direct versus averted eye-gaze in dynamic faces. Eyes moving towards or away from the infant have a very different social meaning, and this lack of discrimination may reflect a marker within the first year of life for difficulties with social understanding. With the same cohort of infants, Elsabbagh, Bedford et al. (under review) used eye tracking to look at scanning of dynamic faces in conditions with different facial features moving alone and together in a peek-a-boo sequence. No differences were found, at 7 or 13 months, in the ratio of looking time to

different regions of the face (i.e., eyes versus mouth) in the high-risk infants, even those who later developed an ASD.

Taken together these findings indicate that it is not a motivation problem early on, as argued by Chevallier et al., (2012) but a specific difficulty in the social understanding. The findings of Chapters 3 and 4 provide evidence for a persistent difficulty in social understanding in high-risk children, particularly those who later develop ASD. Further, reduced looking time and failure to learn from feedback contribute towards reduced receptive vocabulary in these children. However, given that ASD is defined by a triad of impairment, it is also important to understand the role of non-social factors. The following chapter (Chapter 5) addresses the question of how social and non-social behaviours interact with one another over development.

Summary of Chapter 4

- High-risk 24-month-olds, like typically developing toddlers, are able to use the mutual exclusivity principle to fast map pseudo-words to novel objects.
- Low-risk controls can use feedback on their choice to retain the word-object mapping in long-term memory.
- High-risk toddlers do not use feedback, particularly corrective feedback, to learn words, and they perform at chance level when tested for retention of the word-object mapping after a 5 minute delay.
- This difficulty learning from feedback relates to lower receptive vocabularies in high-risk infants.
- Looking time at 13 months from the gaze following task, a measure of referential understanding, is related to word learning performance following reinforcing feedback upon an initially correct object choice. This suggests that there is some continuity in the understanding of social cues over development.
- No significant relationship was found between 13 month looking time and fast mapping using ME, supporting the idea that ME is a learnt heuristic which does not involve social understanding.

Chapter 5

Developmental Interactions between Social and Non-Social Attention in Infants at High Risk for Autism Spectrum Disorder

5.1 Introduction

In Chapters 3 and 4 the prediction of later outcomes, both clinical (i.e., ASD diagnosis) and general (vocabulary and word learning), by early gaze following behaviour were examined. However, ASD is defined by a triad of behaviours which includes both social and non-social difficulties. Thus, looking purely at early social markers, such as gaze following, may be problematic when thinking about the development of ASD. As discussed in Chapter 1, while there have been studies looking at links between non-social executive function behaviours and social cognitive factors such as theory of mind in children with ASD (e.g., Pellicano, 2007; Russell, Saltmarsh & Hill, 1999), little is known about how social and non-social factors might interact over the course of development. Indeed, the majority of etiological theories have focused on a single causal factor.

The first part of this chapter focuses on how a measure of social (gaze following behaviour) and non-social (visual disengagement) attention together predict ASD outcome and symptomatology. This is the first study to look at the effect of more than one measure in relation to the development of ASD, and thus offers an important conceptual step forward in thinking about the potential underlying mechanisms involved. The second part of the chapter addresses the interactive nature of developmental pathways by examining how gaze following and disengagement behaviours relate to one another across the first year of development. Looking longitudinally at the interrelationships between these behaviours over time is important for understanding the developmental route to ASD outcome.

*Part 1: Gaze Following and Visual Disengagement Independently Predict the
Development of Autism Spectrum Disorder*

5.2 The autistic triad

As discussed in Chapter 1, ASD is defined behaviourally by a triad of impairment with difficulties in both social (social and communication difficulties) and non-social functioning (restricted and repetitive behaviours). The majority of cognitive theories of ASD propose a central, causal impairment in either social or non-social skills, which lead to the triad of impairment. However, the attempts to link any single deficit with the whole range of behavioural symptoms seen in ASD have been somewhat unsuccessful (see Chapter 1). It may be that the search for a unitary account has been hindered by the heterogeneity associated with ASD symptomatology. Indeed, Happé and Ronald (2008) argue that searching for a single underlying cause is the wrong approach, because the triad of impairment which defines ASD is in fact ‘fractionable’, with different symptom clusters having different etiologies.

The idea that ASD may have multiple causes has been around for several decades, with Wing and Wing (1971) arguing that ASD should be conceptualised in terms of a combination of impairments. However, Happé and Ronald’s (2008) ‘fractionable triad hypothesis’ formalised this theory on the basis of evidence from both the typical population and individuals with ASD. For example, in children aged 7 and 9 years, population based twin studies have demonstrated only fairly low correlations between autistic behavioural traits in the three cores domains (Ronald, Happé & Plomin, 2005; Ronald et al., 2006), suggesting that, at this age, the behaviours in the triad remain relatively independent.

Factor analysis also offers a useful methodology for addressing the question of whether the triad symptoms are associated. However, the evidence is somewhat mixed. Several studies have presented a single principle component to explain symptomatology (i.e., Constantino et al. 2004; Szatmari et al. 2002; Volkmar et al. 1988; Wadden, Bryson & Ridger, 1991). The existence of one factor explaining the majority of the variance is in line with the idea of a unitary account of ASD. However, Happé and Ronald (2008) point out that 1) of these studies, only Constantino et al. (2004) used item level scores (rather than subscales); and 2) the variance explained by the principle component is not consistently reported, so comparison across the studies is difficult. Further, several studies using factor analysis have found evidence for between three and six factors, indicating a degree of separability among symptoms (e.g., Berument, Rutter, Lord, Pickles & Bailey, 1999; DiLalla & Rogers, 1994; Lecavalier, 2005; Miranda-Linne & Melin, 2002; Stella, Mundy & Tuchman, 1999; Tadevosyan-Leyfer et al., 2003; van Lang et al., 2006; Wadden et al. 1991).

While factor analysis and twin studies provide some evidence for the idea of a fractionable triad, these studies come from children or adults who have already developed the neural systems for processing social and non-social stimuli. The genetic studies have the advantage that, although there are epigenetic changes to DNA over the course of development, broadly, they give a causal account. The existence of genetic differences in symptom clusters thus provides good evidence for multiple difficulties in separate underlying mechanisms. However, in order to fully understand the relationship between social and non-social behaviours as they are emerging and their link with later ASD outcomes, it is necessary to take a developmental approach. Measuring these abilities in high-risk children during infancy, before the onset of overt behavioural symptoms, would further our understanding of the temporal ordering of events and thus get closer to addressing questions of causality.

5.2.1 Cumulative and cascading models of risk

Elsabbagh and Johnson (2010) formalised three different possible models for how risk markers early in development might be related to the emergence of ASD (see Figure 5.1, from Elsabbagh and Johnson, 2010).

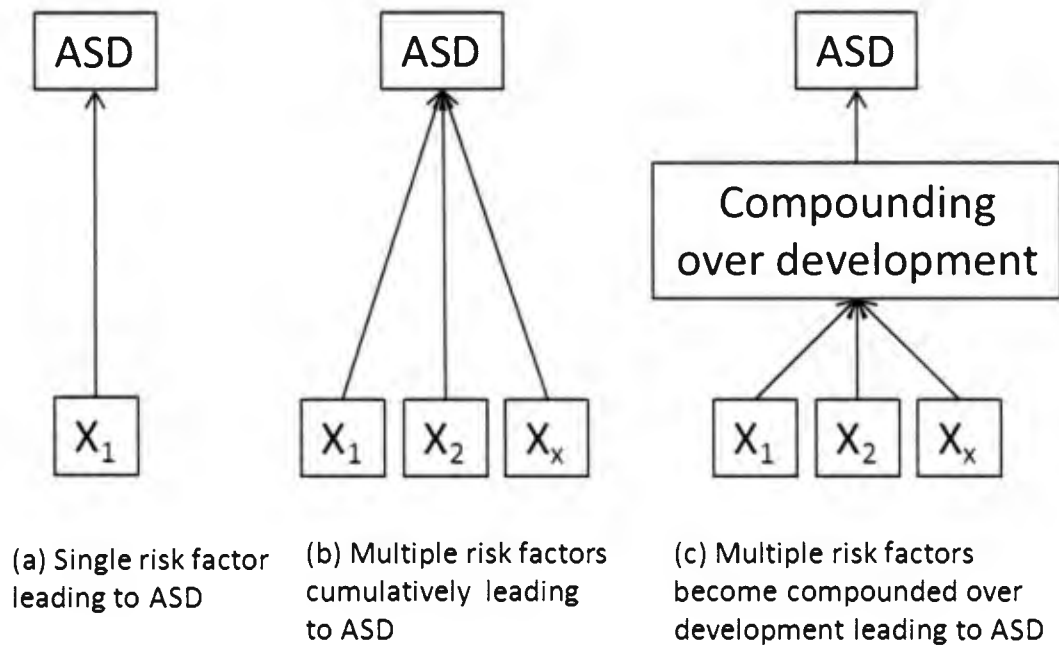


Figure 5.1 From Elsabbagh and Johnson (2010). Three different models for how risk factors might lead to ASD.

The first (a), and most simple, is that a single risk marker results in ASD outcome. This model might suggest, for example, that impairments in understanding social cues are the primary risk factor for developing ASD. However, impairments in non-social attentional disengagement have also been found to contribute to symptoms of ASD (Elsabbagh, Fernandes et al., under review; Elsabbagh et al., 2009a; Zwaigenbaum et al., 2005). Thus we can immediately rule out this single factor causal model as an accurate account of the emergence of ASD.

The other two theoretical models both emphasise a role for multiple risk factors. In the cumulative risk model (b), multiple early risk factors that exceed some threshold work in a

broadly additive way to alter a child's developmental trajectory, resulting in the development of ASD. In the final (c) cascading effects model, on the other hand, the variability in trajectories results from interactions between genetic and environmental factors during development, while the brain is still highly plastic, leading to a non-linear profile of impairment and mapping from risk factors to outcome.

Cumulative and cascading effects risk models make specific predictions in terms of the modularity or interaction between the brain systems subserving social and non-social cognition. Cumulative effects models predict summation, but little interaction between these systems over development, whereas cascading effects models predict the opposite pattern, with development playing a central role in the specialisation of these regions. While these models have been discussed with respect to atypical development (e.g., gene-environment models in Williams Syndrome; Karmiloff-Smith, 2006), there has been no formal empirical testing of these contrasting hypotheses.

This study examines the relationship between gaze following (a measure of social attention) and visual disengagement (non-social attention) with subsequent clinical ASD outcomes, enabling cumulative versus cascading risk models to be tested using developmental data. From previous work it is already known that, independently, both of the measures here (looking time from the gaze following and visual disengagement in the gap-overlap task) relate to ASD outcome (Chapter 3; Elsabbagh, Fernandes et al., under review). However, in this study the combined contribution of these risk factors and their interaction is assessed. Before discussing the specific hypotheses, I will first introduce the literature on visual attention using the gap-overlap task in children with ASD and in high-risk infants.

5.3 Why might visual attention be important in the development of ASD?

Tsotsos et al. (1995) proposed that the human visual attention system is composed of several subcomponents, which include selecting a region or particular features of interest, controlling the flow of information through the visual system and shifting attention from one area to another. The focus of this chapter will be on the disengagement and shifting of visual attention.

Attentional disengagement has been proposed to have an impact on a range of different behaviours. In typical development, visual attention plays an important role in the regulation of emotional states (Johnson, 1990; McConnell & Bryson, 2005; Rothbart, Posner & Rosicky, 1994). Van der Geest et al. (2001) suggest that atypical visual attention may underlie the unusual processing of reciprocal gaze in ASD and Landry and Bryson (2004) further propose that difficulties with both executive function and social orienting stem from early attentional disengagement impairments. They argue that early difficulties in attentional flexibility may contribute to 'restricted temperamental styles' as well as the social and communication impairments that typify ASD.

5.3.1 Attentional disengagement difficulties in ASD

Autistic adults show problems with disengagement on tasks that require rapid attentional shifting to different spatial locations (Casey et al., 1993). More recently, focus has turned to whether early problems disengaging visual attention may offer a developmental explanation for some of the later emerging symptomatology in children with ASD. Several studies have employed the gap-overlap paradigm to investigate visual attention abilities in children with ASD. This task involves presentation of a central stimulus followed by the appearance of a peripheral stimulus in three possible trial types: 1) baseline in which the central stimulus offset is immediately followed by the target peripheral stimulus onset; 2) overlap trials in which the central stimulus remains on the screen whilst the target is presented; and 3) gap

trials where there is a temporal gap between offset of the central stimulus and onset of the peripheral one.

Using this gap-overlap task, Landry and Bryson found that for 20% of the trials autistic children remained ‘stuck’ on the central stimulus for the whole 8 second trial and 80% of children failed to disengage on at least one of the trials. To further investigate this disengagement problem they grouped reaction times (RTs) by speed and found that there were fewer shifts from autistic children in the categories with the shortest RTs. So not only did the autistic children show problems disengaging, they also had particular problems with rapid shifts.

Chawarska, Volkmar and Klin (2010) also used the gap-overlap task in young children with ASD. However, unlike in Landry and Bryson’s (2004) study, in which the central stimuli were blocks ‘falling on each other’, Chawarska et al. (2010) used a centrally presented face and target stimuli (e.g., star shape) to the right and left. Contrary to Landry and Bryson (2004) they found that typically developing and developmentally delayed children failed to disengage on more trials than those in the ASD group. These results argue for *enhanced* attentional disengagement in the ASD group, perhaps suggesting that for autistic children, face stimuli are less engaging than the dynamic blocks used in Landry and Bryson’s (2004) study. Indeed given the evidence that children with ASD show a preference for contingent stimuli (Klin, Lin, Gorrindo, Ramsay & Jones, 2009), the blocks ‘falling on each other’ may have been particularly salient to the ASD group. The slower RTs to disengage seen in typical development may reflect the fact that faces are important and socially relevant stimuli, thus they may be processed more deeply than other types of stimuli.

5.3.2 Attentional disengagement difficulties in infants at high risk for ASD.

In another version of the gap-overlap task, Elsabbagh et al. (2009a) looked at disengagement and facilitation in 9- to 10-month-old high-risk infant siblings of autistic

children. As in Landry and Bryson (2004), the disengagement effect was measured as the difference in RT between baseline and overlap trials. The facilitation effect, on the other hand refers to attentional cueing of a temporal gap before the peripheral target appears. The high-risk infants showed reduced disengagement, reflected in longer RT latency, and less facilitation effect of the preceding gap. While, in this particular study there were no ASD outcome data, the authors do suggest a possible mechanism for the development of autistic symptoms. They argue that difficulty switching attention may result in ‘locking’ onto particular irrelevant parts of a scene. Problems with attentional shifting could enhance, and potentially play a causal role in, difficulties with social orienting (e.g., Osterling and Dawson, 1994). Elsabbagh et al. (2009a) also note that early visual attention difficulties could lead to problems in global processing postulated by weak central coherence theory (Frith and Happé, 1994) as flexible visual scanning enhances global processing of form.

The central stimuli in this study were a cartoon sun and clown, and so fall somewhere in-between the stimuli used by the Landry and Bryson (2004) and Chawarska et al. (2010) in the studies discussed above. The mixed findings across different studies emphasise the importance of thinking about the exact stimuli used. In Elsabbagh et al.’s (2009a) study, the cartoon faces rotated and made an attention-getting noise, potentially representing a more interesting stimulus to high-risk infants than real-life faces. This would explain the similarity between this study and Landry and Bryson’s (2004) findings with young autistic children.

A separate study using the same task with a different cohort of children (the cohort of children used in this thesis) found that although 7 month disengagement (overlap RT – baseline RT) did not distinguish the children who went on to develop ASD, by 14 months the ASD group showed significantly slower disengagement in comparison to the low-risk controls, the typically developing and the atypically developing high-risk infants (Elsabbagh, Fernandes et al., under review). Further, the change from 7–14 months distinguished both

those infants who were atypically developing and the ASD group from the low-risk controls and typically developing high-risk infants. This finding is similar to Zwaigenbaum et al. (2005) who found that a decrease in disengagement ability between 6–12 months predicted ASD as measured by the ADOS-G at 24 months. Twenty-five percent of their high-risk sample showed this pattern of increased latency to disengage at 12 months of age compared with performance at the earlier 6 month assessment. *All* of the infants in this subgroup went on to receive a diagnosis of ASD on the ADOS-G. Zwaigenbaum et al. (2005) also found that measures of disengagement on the Autism Observation Scale for Infants (AOSI) at 12 months were among the risk markers that predicted ASD as measured by the ADOS-G at 24 months.

From Elsabbagh, Fernandes et al.'s (under review) study, based on the same cohort of children as this thesis, it is already known that disengagement in the gap-overlap task, in 14- but not 7-month-olds, relates to ASD outcome. The age at which predictive differences emerge in disengagement is similar to the results of the gaze following study in Chapter 3, in which only looking time to the congruent object at 13 month predicts later symptoms of ASD. Thus, based on these previous studies (see Chapter 3; Elsabbagh, Fernandes et al., under review) we know that both disengagement latency in the gap-overlap task, and looking time in the gaze following task at 13 months relate to subsequent clinical outcomes.

In the current study, I firstly aimed to test whether both of these 13 month measures predict ASD outcome when entered into the same regression model. If the two measures remain independent predictors, this provides evidence against the idea that ASD emerges as the consequence of a single deficit in any one domain (i.e., social or non-social). Independent social and non-social predictors would be consistent with either the fractionable triad hypothesis, which proposes separate social and non-social symptom clusters with different etiologies, or with a cumulative risk model in which multiple factors contribute to ASD

outcome. A significant interaction, on the other hand would imply that the relationship between social processing (looking time) and ASD outcome varies as a function of attentional flexibility (disengagement time), rather than being additive in nature.

Secondly, I aimed to look at the relationship between 13 month disengagement and looking time with the overall growth rate of the Mullen Scales of Early Learning (*MSEL*) Early Learning Composite (ELC) scores from 7–36 months. If disengagement and/or looking time are significant predictors of the rate of growth in ELC scores, this suggests that the measures are not specifically related to ASD outcomes, but predict cognitive development more generally.

5.4 Method

5.4.1 Gap-overlap task: 7 and 13 months

One hundred and four infants (see Tables 2.1 & 2.2) took part in the gap-overlap task at the 7 and 13 month visits (the same age at which they completed the gaze following task). At 7 months, 4 infants (2 high-risk) were excluded and at 13 months, 6 infants (2 high-risk) were excluded owing to non-compliance or technical difficulties.

Infants were seated 60cm from the computer screen on their parent's lap and the stimuli were presented on a 46" LCD monitor. Looking behaviour was recorded using a video camera (see Figure 5.2). The rate of trial presentation was controlled by the experimenter. Each trial began with a centrally presented animation subtending $13.8^{\circ} \times 18^{\circ}$, which rotated to attract the infant's attention. Once the infant looked to the centre of the screen a peripheral target, subtending $6.3^{\circ} \times 6.3^{\circ}$, appeared randomly on either the left or right of the screen (eccentricity of 15°). The peripheral targets were green balloons that expanded and contracted to get the infant's attention. The target remained on the screen until either 1) the infant looked

at it; or 2) the maximum time limit of 2.5 seconds passed. When one of these criteria had been fulfilled, an animal (elephant, lion, seal etc.) appeared accompanied by a sound, replacing the target green balloon.



Figure 5.2 A 13-month-old infant watching an overlap trial in the gap-overlap task.

Infants were presented with a maximum of 70 trials, but the task was stopped sooner if they became fussy. There were two different trial types: baseline and overlap. In the baseline condition, the central stimulus disappeared at the *same time* as the peripheral target appeared, whereas in the overlap condition, the central stimulus remained present (but not moving) while the target stimulus appeared in the periphery. The conditions were presented pseudo-randomly across two blocks (one with a sun and the other with a clown as the central stimulus).

5.4.2 Analysis and measures

In the gap-overlap task, data were video coded frame-by-frame by two raters, who established a reliability of >0.9 (Cohen's K) for trial validity and a correlation of 0.87 for saccadic reaction time on a practice data set. There was no difference in the number of valid trials completed between high- and low-risk infants or between outcome groups (low-risk

control, TD-, AT- ASD- sibs), all $p > 0.12$. Invalid trials were those in which the infant 1) looked away from the screen; 2) did not look at the central stimulus immediately before the onset of the peripheral stimulus; 3) looked away or blinked during the onset of the peripheral stimulus.

Disengagement: Saccadic reaction times were analysed from all valid trials in which the infants oriented towards the peripheral stimulus 100ms–1200ms after its onset. If the infant did not look towards the stimulus in this time the trial was called ‘failure to disengage’ and reaction time was not analysed. These trials were not analysed for the purpose of this thesis. A measure of ‘disengagement’ was calculated as reaction time in overlap trials – baseline trials. This measure of visual disengagement was chosen because, as discussed earlier, disengagement of attention has been argued to underlie the development of gaze following behaviour. See Appendix 5, Figure A5.3 for the distribution of disengagement reaction times, and Table 5.1 for descriptive statistics.

Table 5.1 Descriptive Statistics for Disengagement Scores at 7 and 13 months.

	Low-risk controls		High-risk infants		‘TD-sibs’		‘AT-sibs’		‘ASD-sibs’	
	M	S.E	M	S.E	M	S.E	M	S.E	M	S.E
7 month	n = 48		n = 52		n = 23		n = 12		n = 16	
<i>Disengagement</i>	174.5	17.1	191.0	17.0	182.2	27.0	192.1	36.8	193.7	28.4
13 month	n = 46		n = 52		n = 22		n = 12		n = 17	
<i>Disengagement</i>	138.2	15.6	179.5	21.2	137.3	29.0	136.4	39.4	268.8	38.2

Gaze following measures:

Looking time: This measure is identical to the one presented in Chapter 3: the proportion of looking time to the congruent object on trials in which first look was correct.

First look: This measure is a difference score between the proportion (out of total valid trials) of first looks to the congruent object versus first looks to the incongruent object.

Mullen Scales of Early Learning (MSEL): The Early Learning Composite (ELC), which is a standardised score derived from four of the *MSEL* subscales (visual reception, fine motor, receptive language and expressive language), from all four visits, was used as a measure of general development.

5.5 Results

5.5.1 Prediction of clinical outcome by 13 month measures

As we know from previous work (Chapter 3; Elsabbagh, Fernandes et al., under review) both 13 month looking time to the congruent object and disengagement separately relate to clinical outcome. To confirm this, I first ran two separate multinomial logistic regressions in STATA, one with looking time and the other with disengagement. Both disengagement and looking time scores were first centred by subtracting the mean (of all participants) from each participant's score, and these centred variables were used in the following regression analyses.

The low-risk controls were the baseline group and were compared to the 3 high-risk outcome groups (ASD-sibs, AT-sibs and TD-sibs). Looking time was a significant predictor of the difference between children with ASD versus low-risk controls, $OR < 0.001$, $CI = (< 0.001 - 0.66)$, $p = 0.04$, with low-risk children looking longer to the congruent object. The odds ratio is close to zero because the model is correctly classifying all but one of the controls into the low-risk group. The result is in the expected direction with longer looking

time to the congruent object increasing risk of low-risk classification, although the large confidence interval reflects the difficulties of running this type of analysis with a small sample size. In a separate model disengagement also significantly predicted the difference between ASD and controls, $OR = 1.01$, $CI = 1.00 - 1.02$, $p = 0.003$, with faster disengagement among the low-risk controls. The odds ratio of just over 1 is due to measurement in milliseconds.

Both looking time and disengagement, and the interaction between them, were then included in the same model as predictors of clinical outcome. The interaction term was non-significant ($p > 0.8$ for all contrasts) and was thus dropped from the model, which was re-run with only 13 month looking time and disengagement. Disengagement reaction time remained a significant predictor of the difference between children with ASD versus low-risk controls, $OR = 1.01$, $CI = 1.00 - 1.02$, $p = 0.026$, and looking time became non-significant, although the trend was in the expected direction, $OR < 0.001$, $CI = (< 0.001 - 1.77)$, $p = 0.065$. As expected from the results of Chapter 3, looking time was a marginally significant predictor of the AT-sibs versus low-risk controls, $OR < 0.001$, $CI = (< 0.001 - 3.6)$, $p = 0.088$, but none of the other group contrasts were significant, all $p > 0.56$.

*Table 5.2 Thirteen Month Looking Time * Disengagement Quartiles Cross-tabulation Split by Outcome Group*

High-risk Outcome			Disengagement Quartiles				
			1	2	3	4	Total
Low-risk	Looking Time Quartiles	1	5	5	6	6	22
		2	3	3	5	1	12
		3	0	5	0	2	7
		4	3	2	1	1	7
	Total		11	15	12	10	48
TD-sibs	Looking Time Quartiles	1	4	3	2	6	15
		2	1	2	0	0	3
		3	1	0	2	1	4
		4	1	0	1	0	2
	Total		7	5	5	7	24
AT-sibs	Looking Time Quartiles	1	2	0	2	1	5
		2	1	0	0	1	2
		3	1	0	0	1	2
		4	2	0	1	0	3
	Total		6	0	3	3	12
ASD-sibs	Looking Time Quartiles	1	0	0	4	4	8
		2	0	0	0	1	1
		3	0	3	0	1	4
		4	0	1	0	3	4
	Total		0	4	4	9	17

In order to understand the results of the multinomial logistic regression more fully, a cross-tabulation was created (see Table 5.2), by breaking down disengagement and looking time scores into quartiles, with a value of ‘1’ for the children showing the fastest disengagement and the longest looking time to the congruent object, and ‘4’ representing the quartile with the slowest disengagement RTs and shortest looking time durations. In the ASD group, all of the children fall into the 3rd or 4th quartile for either disengagement or looking

time, with only four children showing scores in the 3rd or 4th quartile for both disengagement and looking time. In other words, it seems that, broadly, looking time and disengagement are predicting separate children with ASD.

5.5.2 Relationship with IQ

Next, the relationship with the *MSEL* composite standard scores was examined, to test whether these task measures relate specifically to later ASD symptomatology, or also to a more global measure of overall development. A growth curve model, using the standardised *MSEL* early learning composite (ELC) scores from all four visits (7, 13, 24 and 36 months) was run (see Figure 5.3). The regression coefficients were fixed at the child's chronological age and an MLR estimator was used. The slope was regressed on 13 month looking time and 13 month disengagement, and the intercept was correlated with looking time and disengagement.

The option for standardised output is not available in Mplus for this type of model so the non-standardised values are reported. There was no significant relationship between the slope of *MSEL* development and 13 month looking time ($\beta = 0.07$, S.E. = 0.05, $p = 0.19$) or disengagement ($\beta = -0.007$, S.E. = 0.005, $p = 0.19$). Nor was either factor significantly correlated with the intercept (looking time: $\beta = 0.001$, S.E. = 0.002, $p = 0.52$; disengagement: $\beta = -0.01$, S.E. = 0.02, $p = 0.50$). There were no significant correlations between the slope and intercept of *MSEL* score ($\beta = -0.001$, S.E. = 0.002, $p = 0.75$), nor between looking time and disengagement ($\beta = -0.01$, S.E. = 0.02, $p = 0.43$).

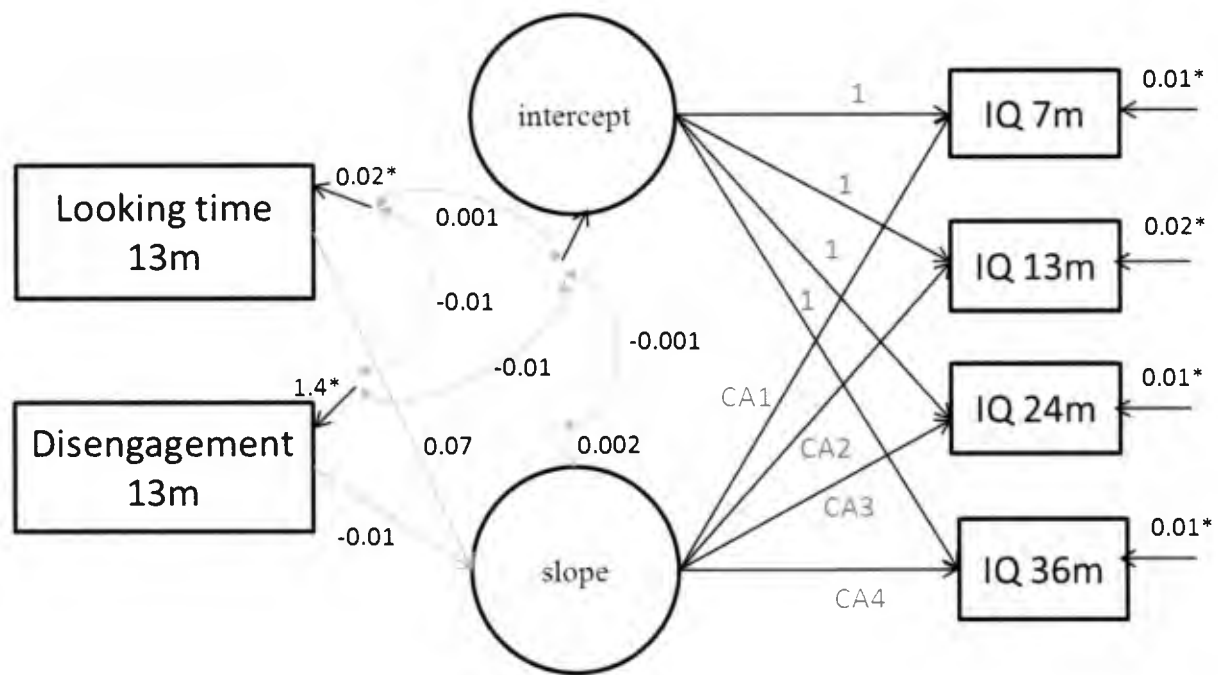


Figure 5.3 Growth curve model to look at the relationship between 13 month looking time and disengagement and the *MSEL* composite scores.

5.6 Discussion

This is the first study to empirically test between unitary versus multiple causal factor accounts of ASD using developmental data. In line with the hypothesis, when included in the same regression model both 13 month disengagement and looking time predicted ASD outcome compared to low-risk controls, although looking time was a slightly weaker predictor, not quite reaching significance. Further, there was no significant effect of the interaction between these measures. These results suggest that a social and a non-social measure of attention both independently contribute to ASD outcome, a finding which is in contrast to the predictions of the majority of theories of ASD (discussed in Chapter 1), which propose a single underlying factor.

The independent prediction of ASD outcome fits with the fractionable triad hypothesis, which argues that social and non-social symptoms are separable (Happé & Ronald, 2008) and thus may have different etiologies. While the independent relationship between looking time

and disengagement with clinical outcome certainly supports the idea of separate risk factors in ASD development, the fractionable triad might go further and predict that early social attention should map more closely onto social-communicative symptoms and non-social attention onto restricted and repetitive behaviours (RRBIs). Thus, we would expect looking time and disengagement to predict different symptoms but in the same individuals.

This is not something which has been directly measured here, and from the results of this current study it is not fully clear whether looking time and disengagement work in an additive way to increase the likelihood of an individual child developing ASD, or whether the measures predict separate individuals who go on to receive an ASD diagnosis. In other words, it is not possible to clearly disentangle whether these two independent measures relate separately to different symptoms or to different cases of ASD. Table 5.2 indicates that the majority of children who go on to receive an ASD diagnosis show difficulties in only one of the two measured behaviours, i.e., in either social or non-social skills, although a small minority do appear to have difficulties in both behaviours, with reduced looking time to the congruent object *and* slower disengagement reaction times. The finding that a small proportion of children have difficulties in both social and non-social abilities is in line with the idea that multiple factors, across separate domains, may be working in an additive way to increase the risk for developing ASD. However, future work with larger sample sizes is needed to test more formally whether fractionated or cumulative or cascading risk models offer the most appropriate description of the data.

The question of individual differences in ASD outcome is important, and not one which the present sample is well suited to addressing, owing to the limited number of children who go on to an ASD diagnosis. It may be that the small subgroup of children who show difficulties in both domains are quantitatively different from those who have impairments in only one domain, i.e., they may have earlier onset ASD, or greater symptom severity. These

are questions which should be addressed by more highly powered studies, as a larger sample size, including many more children who develop ASD, would enable growth mixture models and latent class analyses to be run. These types of models, discussed in Chapter 2, allow different numbers of classes to be fit to the data based on statistics and theory, potentially helping in the identification of different subgroups of children.

A further limitation of the small sample size in the current study is that it pushes the limits of multinomial logistic regression. The odds ratio value for looking time has a huge confidence interval, which is an indication that this analysis should be performed with a larger sample. While Table 5.2 descriptively confirms the results of the regression, i.e., that children with ASD have difficulties with both looking time and disengagement behaviours, future work should aim to combine multiple risk markers in the prediction of clinical outcomes, using larger numbers of children who go on to develop ASD.

Finally, no relationship was found between looking time or disengagement with the rate of global development as measured by the *MSEL*. While there are various studies looking at the links between habituation and later IQ (discussed in Chapter 1), this is the first study to look at how disengagement and gaze following relate to more global development. The fact that no relationship was found indicates that there is a degree of specificity in the relationship between these measures and ASD outcome.

The results of this section suggest that social and non-social attention at 13 months independently contribute to ASD outcomes, with the majority of children who develop ASD showing impairments in only one of either looking time or disengagement. Further, there seems to be a degree of specificity in the prediction of ASD, with the measures showing no relationship with cognitive development more generally. Future studies, using larger samples, should aim to examine individual differences between children who have impairments in one versus two domains at 13 months, and to test whether this distinction is clinically relevant.

*Part 2: Interactions between Gaze Following and Disengagement at 7 and 13 months in
Infants at High Risk for Autism Spectrum Disorder*

5.7 Introduction

This aim of this section is to examine the early relationship between gaze following and disengagement over the first year of life, and to link it with developmental theory. The results presented in Part 1 showed that these measures of social and non-social attention independently predict ASD outcome. However, in order to understand more fully how ASD emerges, it is important to consider the development of looking time and disengagement behaviours over time. The following section addresses the question of how looking time and disengagement become independent predictors of ASD, and 1) whether these abilities are separate and remain separate from early on in development; or 2) whether they become separate over time. These different alternatives have important implications for the way we think about ASD risk factors. If social and non-social behaviours are separate from birth (or soon after) and, as we saw in Part 1, seem to predict different populations of children with ASD, then the implication is that there are at least two independent routes to an ASD diagnosis. In this case, further work should be done into understanding the phenotypic diversity of ASD. If, on the other hand, social and non-social behaviours become increasingly specialised over development, then we should be focusing more on understanding the nature of the changes that take place at a cognitive and neural level during development (possibly through interactions with environmental factors), which bias an individual towards attentional or social difficulties.

5.7.1 Postnatal brain development

The human cerebral cortex is highly specialised for particular functions, with separate cortical regions subserving different skills, such as language or emotion. In comparison to

other mammals, human postnatal brain development occurs over a greatly extended period, which results in those late developing structures having a greater volume (Clancy, Darlington & Finlay, 2000). This protracted development also means that there is an extended period for environmental interactions to shape neural development. There have been different approaches to explaining how these separate and specialised regions emerge through development (Johnson, 2001). Here, I discuss two main theories of postnatal functional brain development, the maturational account and interactive specialisation (IS).

The maturational account is domain specific, proposing that different brain regions reach maturity (the adult state of functioning) and come fully online at different times. This account is in line with modularity theories, which suggest that the processing of particular tasks is subserved by largely separable brain regions. The maturity of a particular region is thought of as a non-linear event which is triggered by genetic or a biochemical change. In many ways this is an intuitive account, but the predictions made by this theory are often not supported by the data. In its strictest sense, the maturational account predicts that behaviours mediated by a particular region should not emerge until that region has reached maturity. However, a recent review of the literature noted that the frontal cortex, which is one of the latest regions to reach neuroanatomical maturity, seems to be involved in various cognitive functions in infancy (see Johnson, 2011). Further, under a maturational account the mastering of a new behaviour should be associated with neural activity in the particular region mediating this skill, but Luna et al. (2001) used functional magnetic resonance imaging (fMRI) to show functional activity in multiple distributed areas during the acquisition of a new behaviour.

One of the alternative accounts, interactive specialisation (IS; Johnson 2001), suggests that the functions of brain regions change and develop through interactions with one another and the environment, a so-called 'activity-dependent' process of change. Initially, cortical regions are poorly differentiated from one another, and multiple regions may be activated for a

particular task. But as these regions interact, activity becomes related to more specific tasks. For example, in face processing the N170/N290 face-sensitive component is modulated by the direction of eye-gaze in typically developing 4-month-olds, but not in young children or adults (Johnson et al., 2005). As discussed in Chapter 1, modulation of this component indicates that the same brain regions are involved in processing both faces and eye-gaze, thus suggesting that there is a developing specialisation of the brain regions involved in face processing. The IS account states that during infancy the social brain (the network of regions specialised in adults for processing social stimuli) is not fully specialised. This may mean that the features of a social stimulus, such as a face, are processed in much the same way as non-social stimuli.

Typically, specialisation in the brain emerges through interactions between brain regions, and environmental input. One potential mechanism by which specialisation is likely to occur is Hebbian firing, in which cells that are active at the same time become associated with one another. If, for example, an infant looks at their mother talking, then cells in visual and auditory cortices as well as ‘social regions’ such as superior temporal sulcus (STS) are likely to be activated. If active cells become ‘wired-together’ then a network of connections across the brain will emerge over time. However, as discussed in Chapter 1, connectivity theory proposes that in ASD there is not a primary deficit in social or non-social processing (or both), but weak long-distance connections within the brain. In the autistic brain, a differing pattern of connectivity could mean that over the course of development different patterns of specialisation emerge.

While models of functional brain development and models of the development of ASD are focussed at different levels, i.e., neural, behavioural and genetic, there are some overarching ideas in common. Development is described broadly as either a modular process, in which separate regions underlie different behaviours, and come online at different times during

development, or an interactive process in which brain regions become increasingly specialised over development through interaction. Under the modular view, the triad of impairment in ASD is likely to be somewhat independent, from the onset, as proposed by the fractionable triad hypothesis, with the disorder arising from multiple hits across separate domains. Interactive theories, on the other hand, stress the importance of developmental interactions among factors over time, which implies potentially multiplicative effects of early risk factors on later ASD outcome.

5.7.2 Links between disengagement of visual attention and gaze following

As discussed in Chapters 1 and 3, various theories have proposed that social attention, such as JA, is particularly, and specifically, impaired in ASD (e.g., Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998). The precise definition of social attention often varies across studies, with the simplest referring to attention within a social context. However, under this view the neural underpinnings should be no different to mechanisms of non-social attention. Social attention has also been conceptualised as processing performed by a dedicated, separate social system within the brain (Johnson, 2001). The work in this thesis goes some way to unravelling these different possibilities, with measures such as ‘first look’ in the gaze following study (Chapter 3) being a social measure only in the widest sense, as it is unlikely to involve interactions among multiple brain regions across the social network. Understanding exactly what is meant by social attention is important when thinking about links between social and non-social measures of attention.

It is plausible that early difficulties with flexibly switching attention lead to problems in social gaze following. In order to engage in joint attention (JA), it is necessary to disengage from a person’s face and switch attention to look at an object. The results in Chapter 3 – that infants who later develop ASD are not impaired in their ability to orient towards the congruent object in response to another person’s gaze shift – argues against the idea that a

simple orienting problem underlies JA difficulties. Further, Leekam, López and Moore (2000) showed a dissociation, with relatively intact disengagement of attention but difficulties in JA, in young preschool children with ASD.

However, in typical development the ability to shift attention flexibly emerges from around 3 months of age (e.g., Hood & Atkinson, 1993). Thus, in order to establish whether disengagement difficulties might be playing some ‘causal’ role in the *development* of gaze following, a prospective design is appealing. The present study aims to examine the concurrent and longitudinal links between a non-social (visual disengagement in the gap-overlap task) and a social (gaze following behaviour) measure of attention in low-risk controls and high-risk infants at two time points, 7 and 13 months.

The first hypothesis was that disengagement and gaze following behaviours will be significantly related at 7 months (i.e., early in development when there has been less time for neural specialisation), with faster disengagement predicting increased looking time to the congruent object and a higher proportion of first looks to the congruent versus incongruent object. However, as social processing becomes mediated by an increasingly specialised network, the strength of these relationships will decrease over the course of development, with the variables less strongly related at 13 months.

Secondly, if disengagement does play a causal role in the development of JA skills, then faster disengagement at 7 months should significantly predict increased gaze following abilities at 13 months, with a higher proportion of congruent versus incongruent object first looks and increased looking time to the congruent object.

5.8 Results

5.8.1 Relationship between visual disengagement and gaze following measures at 7 and 13 months.

An autoregressive cross-lagged model (see Figure 5.4) was run on the whole sample (both high- and low-risk infants) with a maximum likelihood estimator, and provided a good fit to the data ($\chi^2(4) = 2.42$, $p = 0.66$, CFI = 1.00, RMSEA < 0.001).

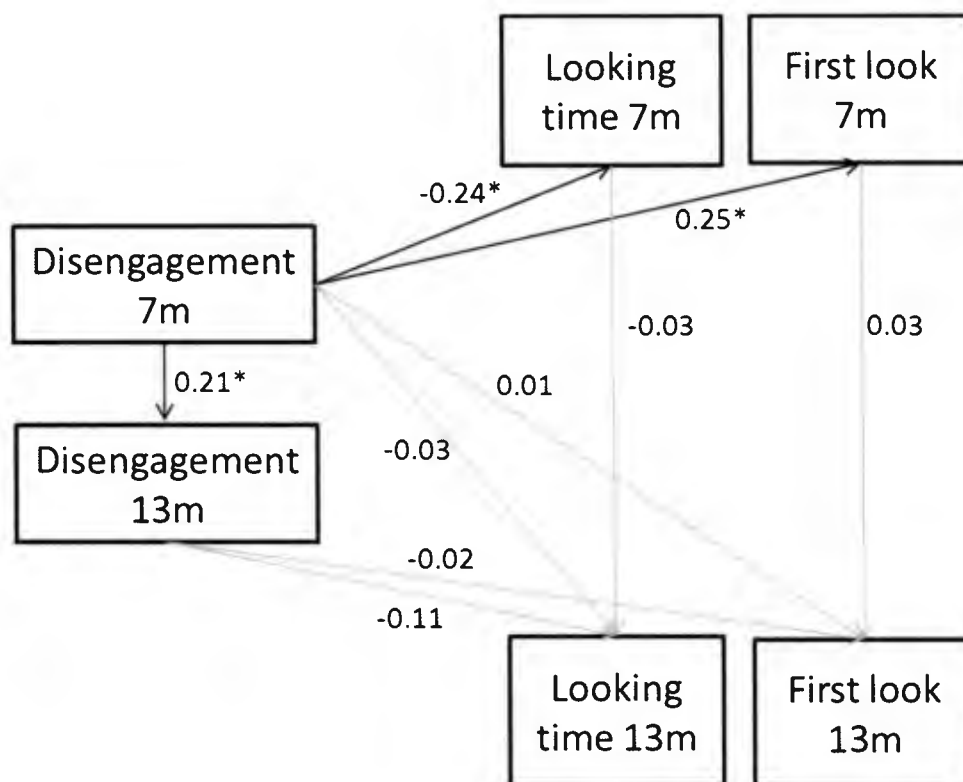


Figure 5.4 Autoregressive cross-lagged model. Correlations among 7 month gaze following measures and among 13 month gaze following measures are not shown in the diagram. The only significant correlation is between 13 month looking time and 13 month first look ($r = -0.24$, $p = 0.04$).

At 7 months, there was a significant relationship between disengagement and looking time ($\beta = -0.24$, S.E. = 0.12, $p = 0.04$) with faster disengagement predicting longer looking time to

the congruent object. The relationship between disengagement and first look difference score (i.e., proportion of correct – proportion of incorrect first looks) was also significant ($\beta = 0.25$, S.E. = 0.11, $p = 0.03$), with *slower* disengagement predicting a higher proportion of first looks to the congruent than the incongruent object (see Figures 5.5a & 5.5b). It is important to note that, although significant, the proportion of variance explained in looking time (0.55%) or first look difference score (0.6%) by visual disengagement is very small. This suggests that while variability in both of the 7 month gaze following measures is predicted by concurrent attentional disengagement, there are many other factors which influence individual differences in gaze following behaviour.

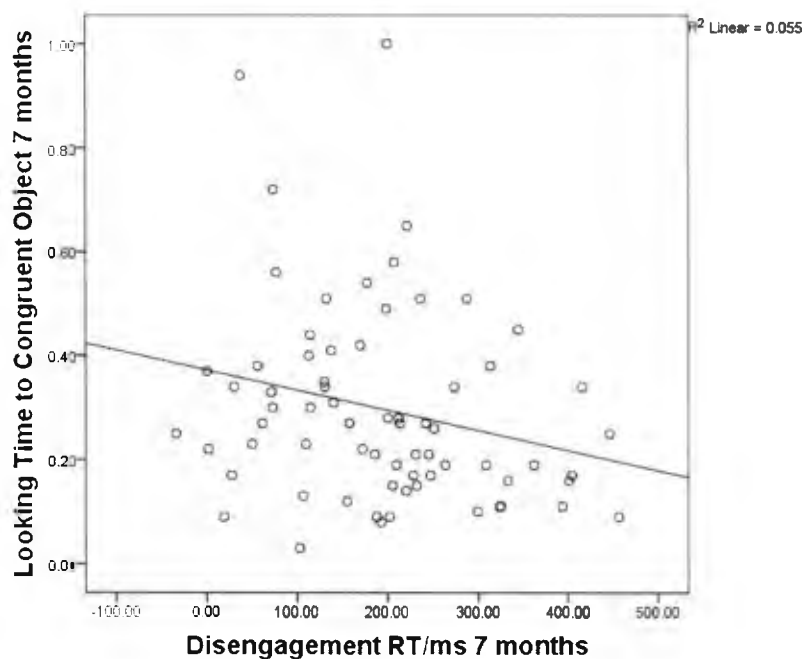


Figure 5.5a Scatterplot showing the relationship between disengagement and looking time to the congruent object at 7 months.

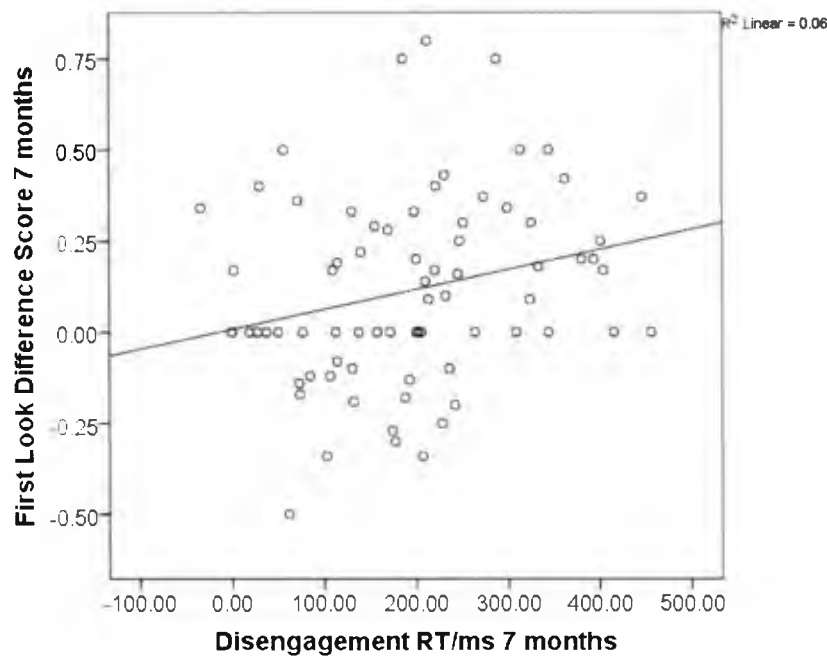


Figure 5.5b Scatterplot showing the relationship between disengagement and first look difference score at 7 months.

Looking at the longitudinal relationships within the same measures over time, disengagement at 7 months significantly predicted disengagement at 13 months ($\beta = 0.21$, S.E. = 0.1, $p = 0.03$). Based on Elsabbagh, Fernandes et al.'s (under review) results, it seems likely that this effect is driven primarily by the low-risk infants, as these children improve in the speed of their disengagement latencies over time, while the reaction times of the atypical and ASD high-risk infants stay the same or actually increase. There were no significant relationships between the 7 and 13 month first look difference scores ($\beta = 0.03$, S.E. = 0.12) or looking time to the congruent object ($\beta = -0.03$, S.E. = 0.12, $p = 0.78$).

Unlike at 7 months, no significant between-measure relationships were found longitudinally for 7 month disengagement and 13 month gaze following behaviours, or concurrently, at the later visit, between 13 month disengagement and 13 month gaze following measures (all $p > 0.36$). Nor were there any significant indirect pathways from 7

month disengagement to 13 month looking time measures via 7 month looking time measures (both $p > 0.78$).

At 13 months, the lack of a correlation in the overall sample between gaze following and disengagement measures could be due to increasing specialisation of the social brain, with disengagement and gaze following becoming processed by separate regions. I wanted to test whether this observed pattern was similar across both risk groups. A significant correlation between the disengagement and gaze following measures at 13 months in the high-risk group could indicate a developmental delay in neural specialisation, with the 13-month-olds in the high-risk group showing similar processing to that seen in 7-month-old infants. However, this was not the case, with neither of the groups showing a significant correlation (both $p > 0.34$).

5.9 Discussion

The hypothesis, that there would be a relationship between disengagement and gaze following behaviour, was supported, but only for the 7 month visit. Faster disengagement at 7 months predicted increased looking time to the congruent object at 7 months in the overall group. One possibility is that those infants better able to disengage from one stimulus and orient to another are quicker to follow eye-gaze and simply have a greater amount of time to look at the congruent object in the gaze following study. However, this seems unlikely given the timescale of only a few hundred milliseconds for attentional disengagement, but a 5 second period in which to look at the congruent object in the gaze following task. It seems more plausible that those infants who are faster at disengaging were able to scan the scene more flexibly and so make more fixations on the congruent object, resulting in an overall longer looking duration. Further analysis, looking in more detail at infants' scan paths, will be required to answer this question more definitively.

The relationship with the first look difference score was in the opposite direction to that predicted, with faster disengagement predicting a greater proportion of first looks to the incongruent than the congruent object. Figure 5.5b shows that children who disengage slowly always have a higher number of first looks to the congruent than the incongruent object. However, of those children who disengage quickly some have a greater number of first looks to the congruent object and others to the incongruent object. It is possible that there is a speed/accuracy trade-off, in which speeded disengagement gives the infant more time to look at the target object, but if disengagement is too fast then the number of orienting errors increases.

It was expected that 7 month disengagement would predict subsequent 13 month gaze following behaviour. However, no significant relationship was found, for either looking time or first looks, which suggests that, contrary to the expectation, disengagement and gaze following behaviours do not interact across development. There was also no significant relationship between 13 month disengagement and 13 month gaze following behaviours. The fact that the 7 month measures relate to one another, but the 13 month measures do not, is in line with the predictions of interactive specialisation (Johnson, 2001). Over development, the increasing specialisation of the social brain network is likely to be important for developing an understanding of the meaning of eye-gaze. Indeed, Pelphrey and Perlman's (2009) model of social brain development predicts that more complex social cognition (i.e., understanding theory of mind) emerges from the increased efficiency of interactions within the social brain network. This specialised processing of social features (as compared to non-social), not present earlier in development, would predict the observed de-correlation between social and non-social tasks over time. Further, the lack of a longitudinal relationship between the same gaze following measures over time could suggest that the measures reflect different underlying processes at 7 and 13 months. In other words, looking time to the congruent

object at 7 months may not yet reflect the level of referential understanding that it does at 13 months. This fits with the emergence of social understanding towards the end of the first year of life (Tomasello et al., 2005), and may also be related to changes in the underlying neural connectivity.

It is likely that there is a developmental change in the neural processes underlying not only gaze following behaviours but also disengagement across the first year of life. Early in development, over the first few months of life, disengagement is driven to a greater extent by bottom-up processing, involving subcortical systems, whereas later in development the frontal and parietal cortices begin to play a role in disengagement processes (Csibra, Tucker & Johnson, 1998; 2001). This top-down control also requires greater communication across a network of brain regions. The fact that the 7 and 13 month disengagement measures were significantly correlated implies that this developmental specialisation may already have occurred by 7 months. However, according to Elsabbagh, Fernandes et al. (under review) the change over time was predominately observed in the low-risk controls and TD-sibs. The AT-sibs and ASD-sibs showed either no change in reaction time or a slight increase, with slower orienting at 13 months. Thus it may be that there is delayed or different neural specialisation in these infants.

The lack of a correlation between 13 month measures was observed in both the low-risk controls and the high-risk infants, which argues against the idea that ASD results from a developmental delay in brain specialisation. However, it may be that in ASD, a different pattern of specialisation is occurring owing to differences in neural connectivity. Future work, should use brain imaging techniques, such as electroencephalography (EEG) and possibly fMRI although this is difficult to use with toddlers because they are required to keep still, to address the neural underpinnings of behaviour.

Another important idea to consider is a limitation of the gap-overlap task methodology, that it is not possible to tease apart engagement from disengagement. In other words, increased reaction time to orient to the peripheral stimulus may be a result of enhanced engagement with the central stimulus, rather than difficulties with disengagement per se. The central stimulus in this task is highly predictable (always a sun in the first block and a clown face in the second block). Holmboe et al. (2010) demonstrated that high-risk infants show increased looking time to a 'boring' repetitive stimulus. While there is an overall significant relationship between disengagement at 7 and 13 months, the limited sample size means separate path analysis models could not be run for each group separately. However, it is possible that at 13 months, the gap-task is measuring increased engagement in the high-risk children who later develop ASD, and this is why no 13 month correlation with gaze following is observed, because the 'central stimulus' in that task is a person's face. As discussed, when a face was used as the central stimulus in the gap-overlap task in young children with ASD (Chawarska et al., 2010) the ASD group actually showed faster attentional disengagement. Future work needs to address the question of whether the disengagement difficulties seen in infants who later develop ASD is stimulus specific, both in terms of the central stimulus, and the peripheral 'reward'.

In conclusion, disengagement of visual attention is related to gaze following behaviour early in development (at 7 months), with speeded disengagement resulting in a higher number of orienting errors when following gaze, but allowing for a greater amount of looking time to the congruent object on the occasions in which orienting is correct. However, as these behaviours develop and become processed by increasingly specialised brain regions, they become de-correlated (by 13 months). Further, no evidence for a longitudinal relationship between early disengagement and later gaze following was found, again suggesting that over development these behaviours may be somewhat separable.

5.10 Conclusion

Gaze following and visual disengagement, measures of social and non-social attention, respectively, separately predict ASD outcome, with high-risk infants who show reduced looking time to the object, and slower disengagement, more likely to develop ASD. This argues against the majority of cognitive theories of ASD, which propose one primary causal factor. The results are more in line with the fractionable triad and cumulative risk accounts of ASD development, although the present study is not able to discriminate clearly between multiple risk factors predicting different symptoms or different cases of ASD. However, it is clear that by 13 months disengagement and looking time have become independent predictors of ASD, and Part 2 of the chapter examined whether these abilities are independent from early on or become independent over the course of development. The findings clearly indicated a relationship between disengagement and gaze following behaviour at 7 months, with the abilities becoming separate over time. This suggests that as the infant develops and the brain networks underlying cognition become increasingly specialised for more complex tasks, behavioural risk markers begin to relate to ASD outcome. In order to understand the emergence of ASD risk factors further, future work should aim to link these observed behavioural changes to the underlying neural mechanisms involved.

Summary of Chapter 5

- At 13 months, a measure of social (looking time to the congruent object in the gaze following task, Chapter 3) and non-social (visual disengagement) attention both independently predict ASD outcome at 36 months.
- The majority of children who develop ASD show difficulties in looking time *or* disengagement at 13 months, with only a small proportion showing difficulties in both abilities.
- The independent prediction of ASD outcome by a social and a non-social measure of attention is in contrast to the predictions of the majority of cognitive theories of ASD, discussed in Chapter 1, which propose a primary causal deficit, and more in line with the fractionable triad account.
- Faster disengagement in the gap-overlap task at 7 months predicts gaze following behaviour at 7 months with longer looking time to the congruent object but also a higher number of orienting errors with more first looks to the incongruent than the congruent object.
- There were no significant relationships between disengagement at 7 months and later gaze following behaviour at 13 months, nor were the measures concurrently related to one another at 13 months.
- The de-correlation of disengagement and gaze following over time is consistent with the idea of increasing neural specialisation, and suggests that future work should focus on the neural underpinnings of these behaviours.

Chapter 6

General Discussion: Multiple Risk Factors in Autism Spectrum Disorder

6.1 Introduction

In this thesis, multivariate modelling techniques were used to examine how early social and non-social behaviours relate both to one another, and to subsequent typical and atypical outcomes. The data came from a longitudinal, prospective study of infants at high risk for an ASD. This design enables the development of ASD to be studied as it emerges and allows us to consider questions of causality. The majority of the theories of ASD specify a single underlying causal factor. However, the emerging findings from prospective studies, together with the results presented in this thesis, suggest that perhaps rather than focusing on a single impairment, we should be looking for multiple early risk markers across domains.

This general discussion chapter will start by summarising the main aims and findings of the thesis. I will then move on to discuss the wider implications of the results, and situate the work presented in this thesis in the context of emerging findings from prospective studies of infants at high risk for ASD. Finally, I will consider the key limitations and ideas for potential future directions.

6.2 Aims

The primary aims of this thesis were 1) to apply multivariate statistical techniques to data from a longitudinal prospective study of infants at high risk for ASD; 2) to investigate how early social behaviours relate to later clinical ASD and language outcomes; and 3) to examine the interaction of social and non-social attention longitudinally over development, and assess their relationship with later ASD outcome (see Figure 1.1 from Chapter 1).

The majority of prospective high-risk studies in ASD have not applied complex statistical methods to analyse their data. This is likely due in part to the limited use of such techniques

across the field of developmental psychology in general. However, structural equation modelling (SEM) offers various advantages over some of the more commonly applied methods of analysis, particularly for longitudinal data. In order to explore the applicability of these methods to the current data set, data from the Mullen Scales of Early Learning (*MSEL*) were used, as this measure provided the most complete data set, available at four time points. Various different models ideal for longitudinal data, including path analysis and growth curve models, were run and provided a good fit to the data.

The types of analysis employed in this thesis are ideal for assessing changing developmental trajectories, and looking at both group and individual level change. In order to understand the variability in development leading to ASD, such techniques will be particularly useful for future research. Given the relatively limited use of more complex statistical methods in the field more generally the widespread application of these approaches is likely to take time, and require substantial work. However, as well as taking account of issues which can bias results, such as missing data, the flexibility of such modelling techniques enables different questions to be addressed. For example, parallel growth curve models can be used to assess links between different but related abilities, and latent class analysis can be used to identify subgroups of children showing different developmental trajectories. Given that ASD can be seen as dimensional with a spectrum of difficulties, but is also defined by a triad of behavioural impairment, flexible models which can include both continuous and categorical variables are likely to be particularly useful here. Further, given that symptoms of ASD emerge over development, modelling the dynamic nature of multiple interacting factors is likely to be an important next step towards understanding the development of the disorder.

Chapter 3 introduced an eye-tracking measure of social attention taken from a gaze following task. Gaze following is thought to be a precursor to joint attention (JA), in which

children share attention with another person about an object of interest. JA type behaviours are among the earliest predictors of later ASD development. In the gaze following study, I aimed to test whether early measures of gaze following, in the first year of life, related to risk status, ASD outcome, and later vocabulary.

Here, I showed that gaze following behaviour can be decomposed into different measures, proportion of first looks to the congruent versus incongruent object, and when the first look was correct, subsequent looking time to the congruent object. These measures are thought to reflect different underlying processes, with first look representing a lower-level orienting bias, and looking time relating to a more complex social understanding of the referential meaning of eye-gaze. Both first look and looking time were measured when infants were 7 and 13 months, but only the latter looking time measure, at 13 months, related to subsequent ASD outcome and atypical development at 3 years of age. Furthermore, this measure predicted later parent reported vocabulary at 24 months, an effect driven predominantly by the low-risk control infants. In line with what we know about links between JA and language, this could suggest a trajectory of social-communicative behaviour across development, manifesting early on in non-verbal responses to social cues, and later in understanding of language.

Chapter 4 investigated another measure of social behaviour, the use of social feedback by 24-month-olds to learn new words. Many children with ASD show early language difficulties, but the potential mechanisms underlying reduced vocabulary are not well understood. High-risk and low-risk children's ability to make a new word-object mapping using the mutual exclusivity (ME) principle, and their memory for the new mapping after a delay was tested. When the child's initial object choice was accompanied by feedback from the experimenter (either reinforcing, if the object choice was correct, or corrective, if the child's choice was initially incorrect) low-risk children were able to retain the mapping in

long-term memory. However, despite being able to correctly apply ME to select the referent of a heard word, high-risk children were not helped by the social cue to learn the word, and performed at chance after a delay. Further, this ability to learn words from feedback was associated with concurrent 24 month receptive vocabulary in the high-risk infants, potentially offering an explanation for the lower vocabularies seen in young children with ASD.

A significant relationship was also found between a measure of referential understanding at 13 months (i.e., looking time to the congruent object in the gaze following task), and children's ability to learn from reinforcing feedback at 24 months, but not their fast-mapping performance using ME. This suggests that the ME principle may reflect a lexical heuristic, rather than involving an understanding of social intent. Further, the relationship between early referential understanding, and later use of social reinforcement to learn words, suggests that there may be a degree of continuity in children's processing of social cues over development.

Chapter 5 addressed the question of how a social and a non-social measure of attention relate to one another, and to ASD outcomes. The fact that the autistic triad includes both social (social interaction and social communication) and non-social (restricted and repetitive behaviours; RRBIs) symptoms suggests that we should be looking at early risk markers across more than one domain. Here, the measure of non-social attention chosen was visual disengagement in the gap-overlap task (the difference in orienting response time between baseline trials, in which the central stimulus is offset at the same time as the peripheral one appears, and overlap trials, in which the central stimulus remains on the screen while the peripheral one is presented). This measure of attentional disengagement was used as a predictor of gaze following behaviour, both infants' orienting in response to another person's gaze shift, and looking time to the congruent object.

Previous work has suggested a link between ASD outcome and both looking time in the gaze following task and disengagement in the gap-overlap task at 13 months (Chapter 3; Elsabbagh, Fernandes et al., under review). In this thesis, I established that when these measures were included in the same regression model both remained independent predictors of ASD outcome (versus low-risk controls) with no significant interaction effect. Further, the different measures tended to predict different children with ASD. This fits with the fractionable triad account, which suggests that social and non-social symptoms of ASD are separable (Happé & Ronald, 2008). However, a small minority of children showed difficulties with both looking time and disengagement, which is consistent with the idea of separate factors working in an additive way to increase a child's risk of developing ASD.

In order to establish whether social and non-social abilities are separate from early on in development longitudinal interactions were examined. At 7 months, the measures were related, with faster disengagement predicting a higher proportion of incongruent rather than congruent first looks, but relating to increased looking time to the congruent object when initial orienting was correct. It may be that there is a speed/accuracy trade off, with faster disengagement resulting in a greater number of orienting errors, but also a longer time to look to the congruent object on the occasions when orienting was correct. No longitudinal relationships were found, and the measures were not related at 13 months, arguing against the idea of early disengagement difficulties leading to later gaze following problems.

6.3 Implications: emergence of behavioural symptoms

The emerging findings from prospective high-risk studies, together with the results presented here, provide almost no evidence for the existence of behavioural markers for ASD within the first year of life (Elsabbagh & Johnson, 2010; Rogers, 2009). However, from 12 months of age, behavioural differences have been consistently reported in infants who later

develop symptoms of ASD (e.g., Zwaigenbaum et al., 2005; Young et al., 2011 etc.). In line with these findings, in this thesis, no risk-group or outcome group differences were found in gaze following behaviour at 7 months, either for first look or looking time. However, by 13 months, infants who later developed social-communication difficulties showed reduced looking time to the congruent object. Further, difficulties in high-risk children's ability to learn words from social feedback were found at 24 months, although this task was not administered earlier on in development.

Further analysis of the relationship between gaze following and word learning revealed that 13 month looking time predicted later word learning following reinforcement. In line with the theory of mind (ToM) account, discussed in Chapter 1, which proposes a social-first deficit in understanding other people's beliefs and desires, this suggests that there might be some continuity in children's use of social cues over time. Both of these measures involve a level of social understanding, and it may be that a difficulty understanding other people's minds is manifested early in development as difficulty in understanding social cues, and later in using such cues to learn. ToM thus offers a straightforward account of this observed continuity in social difficulties across development. The ToM account also argues that problems understanding the social world result in social withdrawal leading to some of the restricted and repetitive behaviours (RRBIs) common in children with ASD. While this question has not been addressed directly here, Happé and Ronald (2008) note that no significant relationships have been found between social functioning and number of RRBIs. Thus, while ToM offers a plausible explanation for some of the observed social problems, it fails to account as well for non-social behaviours.

Returning to the question posed at the end of Chapter 1 of whether social or non-social difficulties emerge first, the findings in this thesis suggest that looking for a primary causal factor in either domain is not the best way to conceptualise ASD development. The results

from Chapter 5 demonstrate that early on in development non-social attentional disengagement and gaze following measures are related, perhaps suggesting that they are initially processed by similar underlying brain regions. However, as the infant develops and the brain becomes increasingly specialised owing to interconnections between different brain regions (Johnson, 2001), as well as environmental input, the social and non-social measures become de-correlated and predict ASD outcome independently.

The fact that multiple measures predict ASD development is perhaps not surprising given that the disorder is defined by a triad of behavioural impairment. The co-occurrence of difficulties across these social and non-social behaviours in individuals with ASD does not necessarily imply a single causal pathway. Further, genetic studies have indicated a degree of separability between social interaction, communication and RRBIs (Happé & Ronald, 2008), suggesting that they may have different genetic etiologies. That these behaviours are somewhat separable poses a problem for the overall explanatory power of ToM. While it offers a good account of the observed social difficulties, it proposes that any non-social problems arise as a secondary consequence of social withdrawal. However, in this thesis using disengagement and gaze following measures as a proxy for non-social and social behaviours, respectively no evidence was found for a primary impairment in social behaviour. Further, no longitudinal relationships between disengagement and gaze following measures from 7 and 13 months were found, suggesting that they are not related over development.

Another difficulty with the ToM account, which also applies to social orienting theories, that argue for a social specific deficit (e.g., Dawson, Meltzoff, Osterling, Rinaldi & Brown, 1998; Dawson et al., 2004; see Chapter 1), is that following eye-gaze, a measure which in many ways can be thought of as 'social', was not impaired in infants who later developed ASD. Defining what is meant by social behaviour is important, because while following eye-

gaze does not require an understanding of intention, it does involve at the very least looking at the person's face (a social stimulus) and responding to its movement. One of the prerequisites for following eye-gaze is looking to the person's face during a period of direct gaze. Indeed, this was one of the inclusion criteria for trials in the gaze following task, with no difference found between groups in the number of valid trials. The fact that high-risk infants attend to the face and correctly follow the gaze shift provides strong evidence against the idea of an overall impairment in the processing of social stimuli from birth.

It could be argued that, although infants at high risk may look at human faces early in development, they are attending to them in a subtly atypical way, by looking only at the mouth, for example. However, no evidence for differences in early face scanning patterns has been found in high-risk infants at 6 or 13 months (Elsabbagh et al., 2012; Elsabbagh, Bedford et al., under review; Young, Merin, Rogers & Ozonoff, 2009). Taken together with the result of intact social orienting, these findings offer strong evidence against social motivation theory (Chevallier et al., 2012). Social motivation theory, discussed earlier in Chapter 1, proposes that ASD development is driven by an early lack of motivation and reward from social stimuli. However, whilst this is a developmental theory, it was not based on developmental data. The findings emerging from the prospective high-risk literature suggest that very early measures of social attention are not impaired in infants who later develop ASD.

As well as finding that high-risk infants show no difficulty with following another person's eye-gaze, in Chapter 4, I also showed that using the ME principle to make a word-object mapping was unimpaired in high-risk children. Both orienting in response to gaze and using a lexical heuristic to map a label to an object are relatively simple processes, which emerge early on in typical development. While these 'lower-level' tasks seem unimpaired in high-risk infants and those who develop symptoms of ASD, there are problems with more

complex social processes which involve understanding the meaning of social cues, such as looking time to the congruent object or using social feedback to learn a word. These more advanced social abilities develop later and are likely to involve coordination across multiple brain regions.

In order to understand why the majority of behavioural difficulties only become observable after 12 months of age, and why only more complex social behaviours seem to be impaired, it is necessary to think about potential underlying neural mechanisms. In the first few weeks of life, neural connectivity in the cortex is based primarily on physical proximity within the brain, and over the following months longer distance connections involving frontal and parietal regions begin to develop (Fransson et al., 2011; Gao et al., 2012). The development of these longer distance connections coincides with increasing top-down control of attention and the emergence of skills such as JA which require integration of information from regions across the social brain network. In order to master more complex social abilities, it is no longer enough only to detect a gaze-shift, but this needs to be linked with other information, such as an understanding of intention.

As discussed in Chapter 1, connectivity theories of ASD argue that the primary deficit is a disruption in neural connections, with weaker long-distance connectivity. Mundy, Card and Fox (2000) found that, compared with responding to JA, initiation of JA requires greater integration of activity between the posterior and anterior attention systems (Mundy et al., 2000; Williams, Waiter, Perra, Perrett & Whiten, 2005). Connectivity theories would thus predict that JA initiated by the child should be more impaired than responding to JA. While both responding and initiating JA difficulties are seen in ASD, initiating difficulties are particularly pronounced. This is reflected in the ADOS-G algorithm score which includes scores from both initiating and responding to JA in module 1 and only initiating JA in module 2 for older children (Lord et al., 2000). This could explain why no early difficulties were

found in following eye-gaze, as this behaviour does not require integration across distal brain regions.

The emerging findings from behavioural and brain imaging studies support the idea of reduced long-distance connectivity in ASD. One of the earliest indicators of neural differences came from the finding that children with ASD have a large head circumference relative to their overall size (e.g., Bailey et al., 1995). Larger head circumference is an indicator of increased brain volume. In young children with ASD (aged 2–4 years) there appears to be an overgrowth in the brain volume which is followed by a drop off with decreased growth (Courchesne et al., 2001). Courchesne (2004) argues that this atypical rate of brain growth may relate to disruption of the formation of neural circuitry, resulting in atypical connectivity.

Wolff et al. (2012) examined connectivity in high-risk infants at 6, 12 and 24 months of age using fractional anisotropy. This gives a measure of diffusion in a particular direction, and thus indirectly reflects axonal myelination and the formation of circuits in the brain. Those infants who showed symptoms of ASD on the ADOS-G at 24 months showed initial overconnectivity at 6 months compared to high-risk infants without symptoms of ASD. However, these effects disappeared at 12 months, and by 24 months the ASD group showed reduced levels of connectivity.

These findings indicate that there are early and persistent differences in the formation of connections within the brain of children who develop ASD. However, Wolff et al. (2012) did not test control, low-risk infants in their study, and so it is difficult to know whether it is the connectivity pattern seen in the ASD group that is atypical, or whether connectivity in the high-risk infants who do not develop ASD actually reflects some kind of protective factor. In other words, it is possible that a certain pattern of connectivity within the brain enables

compensatory neural processing, and thus shifts the developmental trajectory away from ASD development.

These emerging findings of disrupted neural development are in line with the predictions of the connectivity account. Further, atypical connectivity offers a potential explanation for the emergence of behavioural difficulties from around 12 months of age, as well as for the observed impairments across the social and non-social domains. While future work (discussed below) will be needed to test the connectivity hypothesis more fully, the findings presented in this thesis are consistent with the idea that different patterns of connectivity lead to specific difficulties in more complex tasks requiring integration of information across a network of brain regions.

6.4 Alternative Explanations

In Chapter 3, looking time to the congruent object but not orienting in response to another's eye gaze, was proposed to reflect an understanding of the social meaning of eye gaze. Eye gaze is referential in that it refers to a particular object, but understanding the meaning of gaze in terms of the mental state of another person is a more complex social process. Under my definition, looking time is more akin to JA behaviours and involves a level of social understanding. However, historically there has been much debate over the interpretation of gaze following behaviour.

Barresi and Moore (1996) also suggest that the second half of the first year of life marks an important development in infants' understanding of social cues. However, they argue that this develops not due to understanding others' intentionality or having a 'theory of mind', but through intermodal integration of first person and third person representations. Whilst such a matching process likely facilitates the development of social understanding, under this account social understanding itself is not a necessary precursor to JA. I think this is a

plausible mechanism for the development of gaze following behaviour, in which movement of another's eyes and head towards an object is mimicked by the infant creating a mapping between their own and others motion toward the object. However, it is less clear how such a mapping between first and third person representations would relate to differential amounts of looking time to the congruent object.

Carpendale and Lewis (2004) also argue against the theory of mind approach, suggesting that social understanding is constructed within the context of social interaction. They argue against the individualistic approach in which the infant first develops an understanding from their own perspective and then applies this to others. The authors cite Moore and Corkum's (1994) experiment, in which infants initially follow the head turn of another person even when that person's eyes are closed, but later only when their eyes are open, as evidence that social understanding develops through the process of engaging in interactions. Under this view joint attention can occur without social understanding.

As well as the above suggestions that social understanding does not necessarily precede JA, it has also been argued that orienting in response to gaze, the process I argue to be low-level, has also been proposed to play a role in social cognitive processes. For example, McNelis and Boatright-Horowitz (1998) found that low ranking patas monkey direct more eye gaze to high-ranking monkeys, suggesting that gaze in this context has a social meaning.

These are important perspectives to consider, given the argument in this thesis that looking time, but not first look, indexes a level of social understanding. It could be argued that the decreased looking time observed in the infants who later develop symptoms of ASD is due to lower level factors such as attentional processing. In other words, infants who are more distractible will look away from the congruent object more quickly. This interpretation of the results would suggest it is a domain general attention difference that is being measured, rather than a difficulty with social processing per se.

However, there is converging evidence from several sources against this argument. Firstly, no difference was found between groups in the total amount of looking time to the screen, suggesting that it is not an overall problem with attention to the task. Further, the looking time at 13 months does not correlate with attentional disengagement (Chapter 5), suggesting that it is not a problem with flexibility of attention. Secondly, the looking time, but not first look measure relates both to later social communication symptoms in ASD and to word learning using social feedback (Chapter 4). This dissociation provides further evidence that the looking time measure, but not first look, is measuring some level of ‘social’ understanding.

A control task as used by Senju and Csibra (2008) in their original study could also have been included to test for group differences in looking time to the congruent object when the eye gaze was masked (by colourful moving image). This image served to draw infant’s attention (like in the eye gaze condition) but the person’s eye gaze could not be seen. In this non-eye gaze condition, Senju and Csibra (2008) found no gaze following behaviour and no difference in looking time to the congruent versus incongruent object. Thus, I would predict that in this control condition, low-risk infants should show reduced looking to the congruent object in comparison to the eye-contact condition, resulting in similar looking durations across the groups.

6.5 Limitations and future work

Perhaps one of the main limitations with the work presented in this thesis is the modest sample size. The sample of 100 or less participants, including both high- and low-risk infants, limits the complexity of the models that can be run, and prevents the use of a ‘multiple groups’ approach. Multiple group models enable separate models to be run for the low-risk and high-risk groups, which would be particularly useful, as model parameters can then be

constrained to be equal and the fit of the model assessed, essentially testing whether the parameter estimates are similar across the groups. This approach would be more satisfactory than regressing out the effect of group.

The models presented in this thesis were motivated strongly by theory, with clear hypotheses, which means that a relatively small sample size is needed in comparison to more exploratory models. However, in order to know how many participants are required to detect a particular effect, power calculations are useful. In SEM multiple parameters are estimated simultaneously, and so there is no simple way to calculate power for a particular model, as different parameters are estimated with different levels of precision. However, power for target parameters can be calculated. For example, in Figure 2.4a, a test of no direct effects between expressive language at adjacent time points (i.e., fixing the target coefficients at 0) gave a χ^2 difference (in comparison to alternative model) of 67.74 for 3df, with 104 participants. A comparable sample of just 25 would thus be expected to give a χ^2 of 16.28, with a power of 94% using the non-central χ^2 method (see Dunn, Everitt and Pickles, pg. 191). However, the effect in this model is strong and replicated across 3 time points (standardized coefficients of 0.42, 0.43 and 0.52). This power calculation is asymptotic and does not take account of other aspects of model estimation, which in reality would not be possible with such a small sample that is only a little larger than the number of parameters being estimated. In the GCM in Figure 2.5, applying the non-central chi-square method to the Wald test statistics, the test of whether the effect of group in predicting the intercept is zero gives power of 93%, but a power of just 47% for the slope. This suggests that to detect a significant effect for slope here a larger sample is required and emphasises the need for replication of the results.

For future studies attempting to apply SEM to small samples, thought should be given to formulating clear hypothesis driven models. While power can be high for detecting repeated

strong effects, even in a small sample, given the constraints of the model fitting procedure, I would not suggest using a sample of much below 100, with many more participants needed for weaker effects, or more complex models.

A possible extension to the current project would be to increase the available sample size to a minimum of 100 children in each group and, where possible, retest the models, as well as running more complex ones. The data this thesis is based on the 'Phase 1' cohort of children in the British Autism Study of Infant Siblings (BASIS). However, after the start of my PhD, another phase 'Phase 2' of the project has started which includes 75 high-risk and 25 low-risk infants, who are currently being seen for their 36 month visit. While across the phases there are some differences in the tasks administered, some studies, such as the gap-overlap task, have been run with at least 100 high-risk infants (as well as low-risk controls) across a range of ages, including 6, 8, 12, 14, 24 and 36 months. Analysis of this data would allow the types of growth curve models used for the *MSEL* in Chapter 2 to be run, and the sample size would be large enough to think about using a multiple groups approach. However, the modelling would be complicated by the fact that 1) the children come from different cohorts; and 2) different methodology was used (eye-tracking versus video coding) for different cohorts and at different ages. These issues would need to be accounted for in the models, by allowing different measurement errors, for example.

Other limitations relate to the structural equation modelling approach more generally. A structural equation model splits the observed variance-covariance matrix between the relationships specified in the model. Thus if a model is incorrectly specified, with a pathway that should be included actually missed out from the model, then the estimates of the other pathways will be biased. While the models used in this thesis were strongly motivated by theory, as well as providing a good fit to the data, they are only one interpretation of the results, and it is likely that other similar models would also have fit the data. An alternative

framework, which I have not used in this thesis but which might be interesting to explore in the future, is the causal inference approach. When a SEM is correctly specified the parameter estimates will be the same as those generated by a causal inference model. However, the definition of ‘casual’ under the causal inference framework is more explicitly and mathematically defined.

A second main limitation relates to the lack of follow-up data beyond 3 years. The rate of ASD clinical outcomes in the current study (32.1%) is higher than in other published studies (Ozonoff et al., 2011; Rogers, 2009). Further, this recurrence rate also falls above the upper 95% confidence interval of 25.5% reported by Ozonoff et al. (2011). This may simply reflect higher error in a relatively small sample, but caution is needed when considering the current findings until the cohort have been followed-up to an age when diagnosis is considered more stable ~4–5 years (e.g., Charman & Baird, 2002).

A third limitation of the work presented in this thesis is the purely behavioural nature of the measures. While underlying developmental processes and neural functioning can be inferred from behavioural data, ideally, future work should look at other electrophysiological and neuroimaging techniques. For example, the co-registration electroencephalography (EEG) and eye-tracking has the potential to identify brain activity and known neural components associated with cognitive processes during active viewing of naturalistic scenes such as the gaze following videos used in this thesis. While combining EEG and eye-tracking poses various technological challenges, such as the electrical artefacts which are introduced into the EEG signal by the oculomotor muscles during eye movements, new techniques are being devised to overcome these technical limitations and record online brain activity during active vision (e.g., Dimigen, Sommer, Hohlfeld, Jacobs & Kliegl, 2011).

One further limitation which should be considered concerns the nature of infancy data. Babies and toddlers often have only a short attention span, and easily become distracted or

fussy. In the gaze following and word learning studies (Chapters 3 & 4) only a very small number of trials were run with each child. This makes it difficult to reliably look at individual differences. This is due in part to the balance that exists in a prospective study such as this, between collecting enough data on a range of different tasks in a limited time frame.

A final issue to discuss is the fact that researchers conducting the ADOS-G assessments were not blind to group status, and they knew whether children were at high or low risk. However, care was taken to ensure that the team who saw infants for the first two visits, in which the gaze following task and gap-overlap tasks were conducted, were not the same researchers carrying out the ADOS-G assessments at 24 and 36 months, and assessments were double coded with high inter-rater reliability.

6.6 Specific study modifications

So far the ideas discussed for future work have arisen from specific limitations in the design and methodology of this thesis. However, the results from the individual studies presented here also raise further questions for investigation, both in terms of ideas for specific studies as well as for prospective studies more generally.

Perhaps the most important idea for future research arising from the findings in this thesis (together with those emerging across prospective studies of high-risk infants) relates to the fact that no behavioural differences were found within the first year of life, but emerge from around 12 months of age. This suggests that important developmental changes are occurring between 6 and 12 months, and fits with what is known about changes in social cognition in typical development towards the end of the first year of life. The majority of high-risk prospective studies test children at 6 months and again at 12 months, but the findings clearly suggest that assessing infants at an intermediate time point, such as 9 months, will be important for studying these differences as they emerge.

A more general point, following on from this, relates to the motivation behind prospective studies. To-date the focus has been primarily on the search for early markers, with research being guided by what we already know about symptoms of ASD. This is, at least to some extent, the approach taken in this thesis, for example, I looked at gaze following behaviour because it is known that JA problems are characteristic of ASD. However, I have also attempted to go beyond this and look at measures over time, as well as interactions between different measures. There is a clear need for this type of hypothesis-driven approach to prospective high-risk studies, informed by an understanding of both typical and atypical developmental processes. Studies aimed at testing theoretical accounts of ASD development will extend our understanding of the underlying brain mechanisms, and provide targets for future research.

As well as these more overarching ideas for future directions, there were also several more specific ideas for future studies arising from the work in different chapters. In the mutual exclusivity task, the high-risk children did not use social feedback to learn word-object mappings, particularly when their initial choice was incorrect. However, there were two potential explanations for this finding that were discussed in Chapter 4. It is possible that the high-risk children simply do not weigh the information given to them by the examiner ('the social feedback') as strongly as their own internally generated hypothesis about the word-object mapping. However, an equally plausible explanation is that, having chosen the incorrect object, high-risk children persevere in failing to make the correct mapping. As discussed, children with ASD are often described as having a cognitive style which lacks in flexibility.

In order to disentangle these two hypotheses I have designed a further experiment which is currently being run with a different cohort of high-risk 24-month-olds. This new study is in many ways similar to the previous task, but a key manipulation is that it involves both a child

and an experimenter taking turns to “play the game”, and map a pseudo-word to a novel object.

In this study, there are four trials, the order of which is counterbalanced into one of two pre-specified orders. Two of the trials are completed by the child and two by the other experimenter while the child watches. In each trial the child/experimenter is asked to ‘give the moxi’ or another pseudo word. In all trials, both objects are novel and so the ‘correct’ answer depends on the child’s/experimenter’s choice, rather than on the object itself, i.e., in a ‘correct’ trial it is the object the child/experimenter chose, and in an ‘incorrect’ trial it is the other object, the one that they did not choose. This means that all children (and the experimenter) complete an equal number of reinforced and corrected trials (unlike in the mutual exclusivity study, Chapter 4). This study design allows a minimal number of trials to be run, which is important when children are participating in multiple studies within a single testing session. Piloting of this task showed that the 24-month-old children were happy to make an object choice, despite the fact that both objects are novel and the non-word could refer to either object.

After a 5 minute break the child is tested on all four target objects (two from trials in which the child made the choice, and two in which the experimenter made the choice). For all participants there are thus two reinforced (i.e., correct initially) and two corrected (i.e., incorrect initially) trials. Following the results in Chapter 4, the predictions are that for the ‘child choice’ trials, the low-risk children, whether reinforced or corrected, should perform above chance. The high-risk children should not be above chance for either corrected or reinforced, although they should show particular problems with the corrected condition. For the new experimenter condition the hypothesis concerning absolute performance level for the low-risk controls is a little less clear. However, it seems likely that control children will find this condition harder than when they make the object choice themselves. The important

comparison here, however, is how the controls perform *relative* to the high-risk group. The prediction is that the high-risk children, if they are perseverating with incorrect choices will be relatively better in the experimenter condition, as they have not made an object choice themselves. However, if they are weighing their own choice above social feedback, then their performance in experimenter condition should be worse than controls, as only social feedback is available to learn the word.

Another specific idea for future research relates to the contrasting of social and non-social attention. In Chapter 5, the measures chosen were disengagement and gaze following, because disengagement of attention has been suggested as a potential underlying mechanism for the development of gaze following behaviour. However, it would be interesting in future to run a gap-overlap task with more carefully controlled stimuli, manipulating the central and peripheral stimuli and including more ‘social’ as well as non-social stimuli.

Chawarska et al.’s (2010) study demonstrated the importance of the central stimulus. They used human face pictures and, contrary to previous studies which showed slower disengagement in ASD, they found speeded disengagement in young children with ASD. In the gap-overlap task used in this thesis, the central sun and clown stimuli had cartoon faces and rotated and made a noise to gain infant’s attention. While cartoon faces are clearly very different to real life photos of human faces, it is not possible to know whether all infants processed them as non-social. The increased disengagement time in children who later developed ASD (Elsabbagh, Fernandes et al., under review) could suggest that the stimuli are not being treated as ‘social’, but it is difficult to know as a gap-overlap task with human faces has not been run with high-risk infants at this young age.

The predictability of the stimuli used in our study may also have played a role. The stimuli rotated in a predictable manner, as well as being repeated from trial to trial (unlike the changing reward stimuli). It is possible that the children who later developed ASD were

attracted to these highly predictable stimuli and found the ‘reward’ less rewarding, and were thus less motivated to orient away from the central stimulus. It will be important in future work to disentangle the mechanisms underlying engagement and disengagement of attention in these infants, by manipulating the choice of central and peripheral stimuli. Further, in order to understand the relative orienting in response to social and non-social stimuli over time, a gap-overlap task which includes both human faces and non-social stimuli would be helpful. Although the gaze following measure used here was an advantage in that it was more ecologically valid and ‘social’ than a photograph of a face, comparing performance within the same task would allow direct comparison of infants’ responses to social and non-social stimuli.

6.7 Implications for intervention

Although ASD is a developmental condition, at present there is only very limited understanding about the way in which interactions between brain and behaviour characteristics influence children’s developmental trajectories. Understanding how risk factors work might work together over time is necessary for developing earlier screening measures and more efficiently targeted intervention strategies. The findings presented in Chapter 5 of this thesis, that both social and non-social measures of attention (looking time and visual disengagement) independently predict ASD outcome, suggest that early intervention strategies should ideally target multiple domains.

One study has looked at the effects of training attention control on performance in the gap-overlap task in typically developing 11-month-old infants (Wass, Porayska-Pomsta & Johnson, 2011). The study involved training attentional control, using a contingent eye-tracking technique in which infants had to track a moving stimulus with their eyes while ignoring any distractors that appeared. They found that infants who received the training

(which was just 77 minutes on average) showed a significant reduction in their disengagement reaction time in the gap-overlap task. Given the results presented in Chapter 5 of this thesis, showing that slower disengagement at 13 months related to a different group of individuals with ASD than those who showed reduced looking time in the gaze following task, this attentional training could potentially offer a useful intervention.

However, the findings in this thesis suggest that, even if such training were to be successful in reducing disengagement reaction time, a separate intervention would also be required in the social domain. In Chapter 3, I found that looking time to the congruent object at 13 months, which is thought to reflect an infant's understanding of the meaning of eye-gaze, was reduced in those infants who later developed social and communication difficulties. Given that it is looking time and not first look which is predictive of ASD outcome, appropriate intervention strategies are likely to be somewhat similar to current JA interventions.

In a pilot study, Aldred, Green and Adams (2004) targeted shared attention as well as parental sensitivity and responsiveness in their intervention with 2- to 5-year-olds with ASD. The intervention aimed to train parents to adapt their communication to the child's particular level. At post-test they found a significantly greater reduction in ADOS-G scores in the active treatment group compared to control participants (although this effect has not been replicated in a large, randomised control trial, Green et al., 2010) as well as a significant improvement in children's expressive language as indexed by the CDI. Kasari, Paparella, Freeman and Jahromi (2008) found that interventions targeting JA are most effective for children with relatively low levels of language abilities and object exploration. This suggests that JA may be a useful starting point but after this initial 'stage setting' other skills, such as imitation, become important for the continued development of language and social skills (Toth,

Munson, Meltzoff & Dawson, 2006). If this is the case then intervening at younger ages is likely to be particularly beneficial.

Given the high rate of ASD as well as sub-clinical symptoms of ASD in high-risk siblings, developing appropriate early intervention strategies offers the potential for ameliorating later severity of symptoms. While the large randomised control trial (Green et al., 2010) looking at parent mediated intervention showed no significant change in ADOS-G score, this trial was conducted with children who already had a diagnosis of core ASD. It is possible that earlier intervention, before the onset of overt behavioural symptoms, could have a more far-reaching effect.

In order to time intervention strategies to maximise their impact, understanding the developmental process is important. The results in this thesis are in line with the growing literature suggesting that behavioural symptoms first emerge at 12 months of age. Further work is needed to understand the neural underpinnings of this developmental change, in order to know whether this might be a good age at which to intervene. Secondly, the results here suggest that interventions are needed across more than one domain, a finding which is in many ways contrary to the majority of cognitive theories of ASD, which posit a single underlying risk factor, and more in line with the predictions of the fractionable triad.

6.8 Conclusion

In this thesis, a range of different multivariate statistical techniques have been applied to developmental data from infants at familial high risk for developing ASD. Early social difficulties were found in these infants across development, with reduced looking time to the congruent, gazed-at object in a gaze following task at 13 months, and problems using social feedback to learn words at 24 months. However, in line with the emerging findings from

high-risk prospective studies, no behavioural difficulties were observed within the first year of life. This suggests that there are important developmental changes occurring from 6–12 months, and that prospective studies should take a more theoretically motivated, hypothesis-driven approach, and should test children in-between these ages. Here, I tested the interrelationships between a social and non-social measure of attention and their prediction of later ASD outcome. This is the first attempt to combine measures, using multivariate techniques, and to test between different developmental models of ASD. The finding that both disengagement of visual attention and looking time to the congruent object independently predict later ASD diagnosis argues against the majority of cognitive theories which propose a single underlying causal factor.

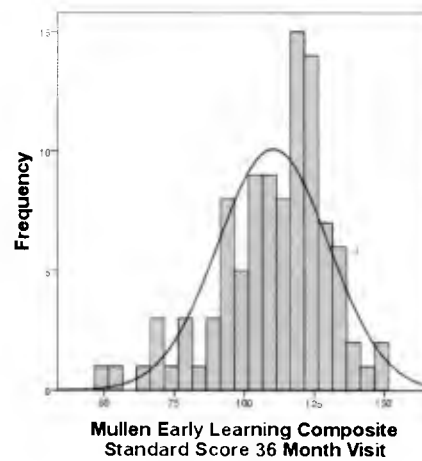
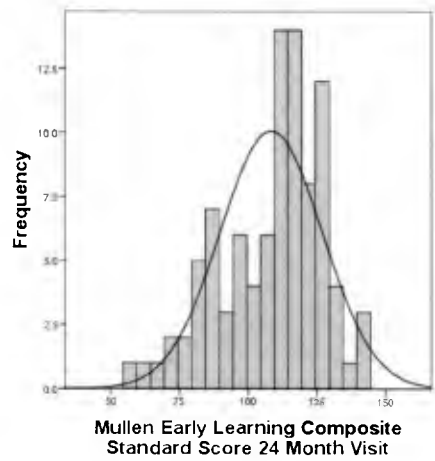
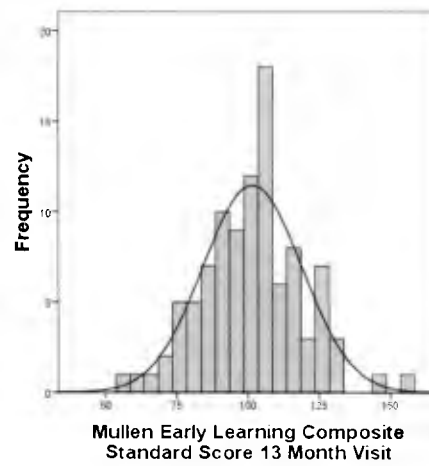
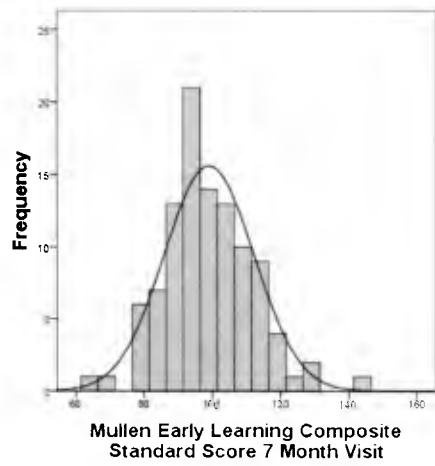
In the future, studies should aim to use larger sample sizes, which will enable the use of more complex statistical techniques, to look at the interactions between multiple measures over time. Further, co-registration between behavioural measures, such as eye tracking and neuroimaging techniques such as EEG, will further our understanding of the neural correlates underlying behaviour, and allow better characterisation of the variable developmental trajectories leading to ASD.

Appendix 1: BASIS Network Protocol Summary

	Measure	Visit 1 (6-9 m)	Visit 2 (12-15 m)	Visit 3 (24 m)	Visit 4 (36 m)
Parent report	Vineland Adaptive Behavior Questionnaire	√	√	√	√
	MacArthur Communicative Development Inventory (CDI)	√	√	√	
	Infant Behaviour Questionnaire (IBQ)/ Early Childhood Behavior Questionnaire (ECBQ)/ Childhood Behavior Questionnaire (CBQ)	√	√	√	√
	Toddler Early Development Inventory Q-CHAT		√	√	
	Social Responsiveness Scale (SRS)				√
	Social Communication Questionnaire (SCQ)				√
	Child Behavior Checklist (CBCL)				√
	Sensory Profile (SP)				√
	Feedback from parents on project	√	√	√	√
Lab (or home visit)	Demographics	√	√(updates)	√(updates)	√(updates)
	Medical history	√	√(updates)	√(updates)	√(updates)
	Mullen Scales of Early Learning	√	√	√	√
	Head circumference & height	√	√	√	√
	Autism Observation Scale for Infants (AOSI)	√	√		
	Autism Diagnostic Observation Schedule (ADOS)			√	√
	Autism Diagnostic Interview (ADI)				√
	Development and Wellbeing Assessment (DAWBA)				
	Parent Child Interaction (PCI)	√	√	√	√
	Development and Wellbeing Assessment (DAWBA)	√			
	Social Communication Questionnaire (SCQ)	√			

Shaded grey areas indicate that the measure will not be collected/is not suitable for administration at that age.

Appendix 2: Distribution of Mullen Early Learning Composite scores at 7-36 month visits



Appendix 3: Mplus model scripts

Figure 2.4a First order autoregressive model for raw expressive language (EL) score at 7 – 36 month visits.

Model:

```
EL_Raw4 on EL_Raw3;  
EL_Raw3 on EL_Raw2;  
EL_Raw2 on EL_Raw1;  
EL_Raw1 EL_Raw2 EL_Raw3 EL_Raw4 on Group;
```

Figure 2.4b Second order autoregressive model for raw EL score at 7 – 36 month visits.

Model:

```
EL_Raw4 on EL_Raw3 EL_Raw2;  
EL_Raw3 on EL_Raw2 EL_Raw1;  
EL_Raw2 on EL_Raw1;  
EL_Raw1 EL_Raw2 EL_Raw3 EL_Raw4 on Group;
```

Figure 2.4c Latent variable first order autoregressive model for raw EL score at 7 – 36 month visits.

Model:

f1 by El_Raw1@1;

f2 by El_Raw2@1;

f3 by El_Raw3@1;

f4 by El_Raw4@1;

f1 with f2@0;

f1 with f3@0;

f1 with f4@0;

f2 with f3@0;

f2 with f4@0;

f3 with f4@0;

El_Raw1 (b) El_Raw2 (b) El_Raw3 (b) El_Raw4 (b);

f4 on f3*1;

f3 on f2*1;

f2 on f1*1;

f1 f2 f3 f4;

f1 on group*-1.3;

Figure 2.5 Example of a GCM using EL score at 7 – 36 months.

Model:

```
i s | EL_Raw1@0 EL_Raw2@1 EL_Raw3@3 EL_Raw4@5;  
i with s;  
i s on group;
```

Figure 2.7 Example of a GCM using chronological age for EL score at 7 – 36 months.

Model:

```
i s | EL_Raw1 EL_Raw2 EL_Raw3 EL_Raw4 at CA_yr1 CA_yr2 CA_yr3 CA_yr4;  
i with s;  
i s on group;
```

Figure 2.9 Four class growth mixture model for *MSEL* expressive language (EL) scores.

Model:

```
%overall%  
i s | EL_Raw1@0 EL_Raw2@1 EL_Raw3@3 EL_Raw4@5;
```

Figure 4.5 Path analysis model looking at the relationship between 13 month looking time and 24 month fast mapping and feedback performance, controlling for risk status.

Model:

```
Fastmap Feedback on LT_13m group;  
LT_13m on group;
```


Figure 4.6 Path diagram of the relationship between 13 month looking time behaviour and mutual exclusivity and later word learning following reinforcement and correction.

Model:

Fastmap Rein Corr on LT_13m group;

LT_13m on group;

Figure 5.3 Growth curve model to look at the relationship between 13 month looking time and disengagement and the *MSEL* composite score.

Model:

i s | elcss1 elcss2 elcss3 elcss4 at CAyr1 CAyr2 CAyr3 CAyr4;

s on Dis_13m LT_13m;

i with s Dis_13m LT_13m;

Appendix 4: Growth Mixture Model

I have used 200 random starts for a two class model and up to 2000 for a four class model, with a sample of ten picked randomly to make sure they converge to the same value. I started by fitting a two class model and increased up to four, a four class model provided the best fit to the data (BIC = 2091, entropy = 8.88), see Figure A3.1. Further, the diagonal of the classification table (Table 2.3) shows that the certainty of being assigned to a given class is very high, in other words there is not much overlap between classes.

Table A3.1 Classification Table for the Four Class Growth Mixture Model

Class	1	2	3	4
1	0.96	0.04	0	0.01
2	0.05	0.94	0	0.01
3	0	0	1	0
4	0.02	0.11	0	0.86

Classes 1 and 2 (red and green lines, Figure 2.9) contain the majority of the children, 32.3% and 58.9% respectively. These classes show the most steeply increasing EL trajectories, with scores in Class 1 starting a little lower than in Class 2, but increasing to a similar point by the final 36 month visit. Classes 1 and 2 contain children from both the high- and low-risk groups, although Class 1 (with the slight lower intercept) contains a higher proportion of high-risk (68.75%) as compared to low-risk children. Class 2 on the other hand contains a higher proportion of low-risk control children (60.32%). Class 3 only contains 3 children, (1 high-risk and 2 low-risk). Thus, despite a four class model providing the best fit, this particular class is difficult to interpret meaningfully.

The final class, Class 4 (black line, Figure 2.9), has a much shallower trajectory than the other classes. This class contains 6 children, all of whom are from the high-risk group. Four of these children were classified into the ASD-sibs group at 3 years of age, and 2 fell into the AT-sibs group. In other words, none of the children in this class, with what appears to be emerging expressive language difficulties, were thought to be typical in their development at 36 months.

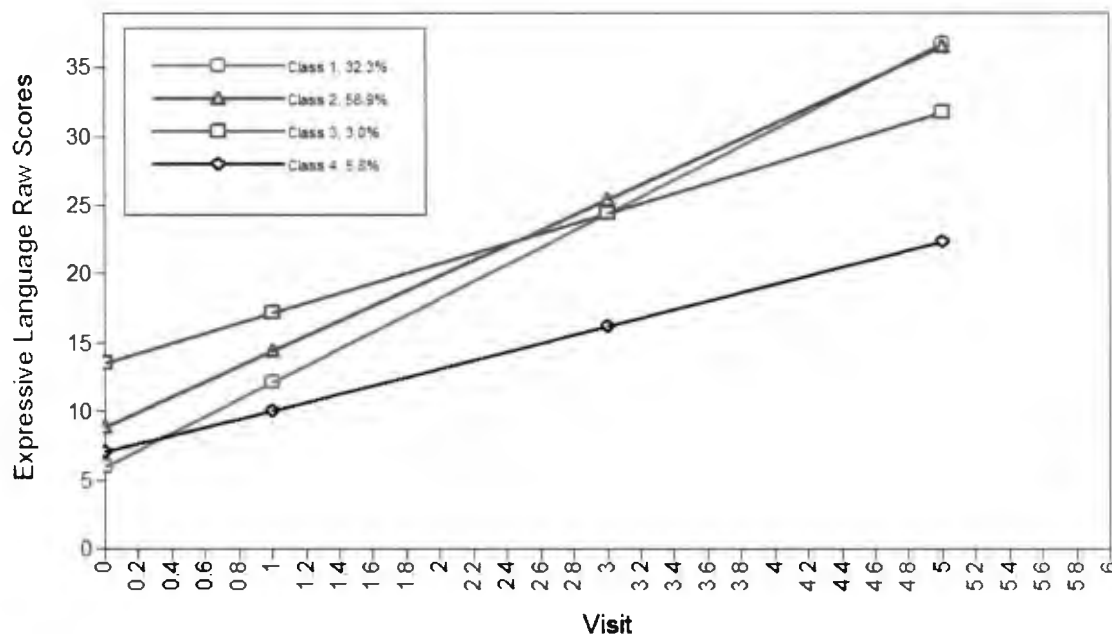


Figure A3.1 Four class growth mixture model for *MSEL* expressive language (EL) scores.

This model suggests that expressive language abilities do not map directly onto four separate clinical outcome classes (low-risk, TD-sibs, AT-sibs and ASD-sibs). However, there may be a subgroup of high-risk infants, who later develop symptoms of ASD, with a reduced rate of growth in expressive language from 7-36 months of age. However, a larger sample size is really needed to retest this model and establish whether there are any clinically significant differences between this group of children with emerging language difficulties, and children with an ASD diagnosis who do not show expressive language problems across the first three years of life.

Appendix 5: Distributions of measures

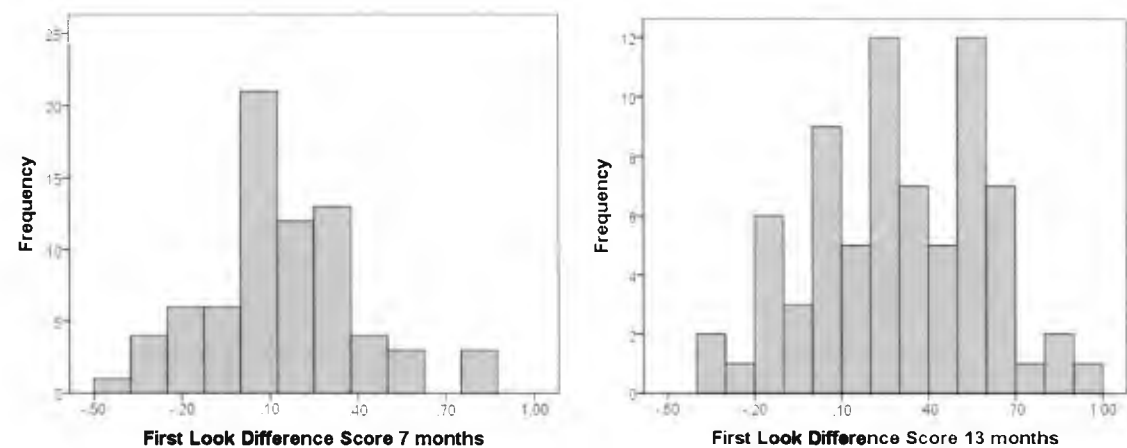


Figure A5.1: Histograms to show the distribution of First Look Difference Scores at 7 and 13 months

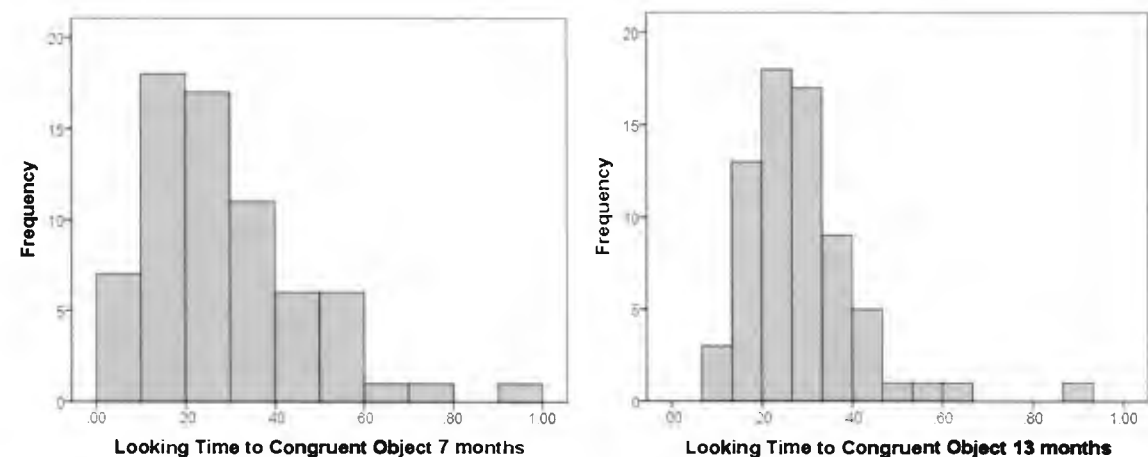


Figure A5.2: Histograms to show the distribution of Looking Time Scores at 7 and 13 months

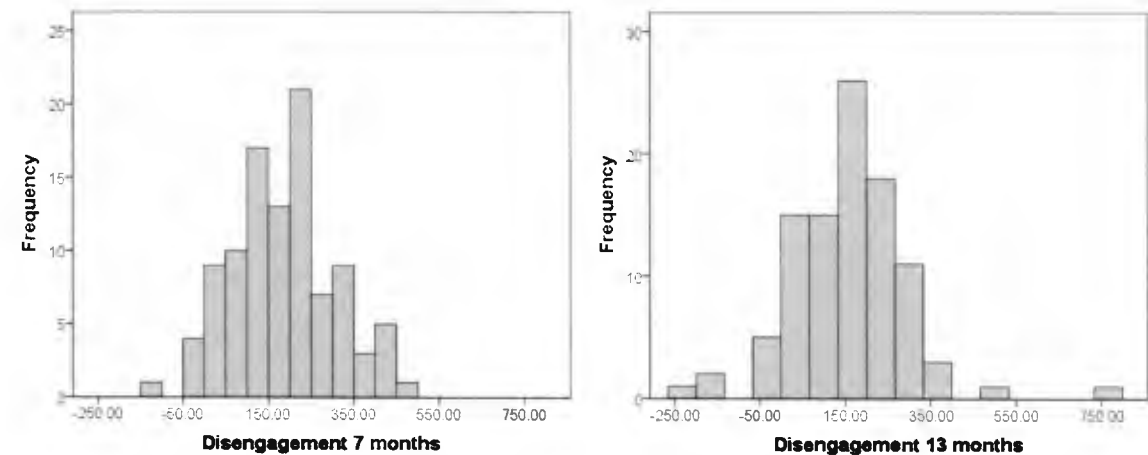


Figure A5.3: Histograms to show the distribution of Disengagement Scores at 7 and 13 months

References

- Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, 4, 165–178.
- Aldred, C., Green, J. and Adams, C. (2004). A new social communication intervention for children with autism: pilot randomised controlled treatment study suggesting effectiveness. *Journal of Child Psychology and Psychiatry*, 45, 1420-1430.
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.
- Anderson, S.W., Bechara, A., Damasio, H., Tranel, D. and Damasio, A.R. (1999). Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Nature Neuroscience*, 2, 469–479.
- Anderson, D. K., Lord, C., Risi, S., DiLavore, P. S., Shulman, C., Thurm, A., Welch, K. and Pickles, A. (2007). Patterns of growth in verbal abilities among children with autism spectrum disorder. *Journal of Consulting and Clinical Psychology*, 75, 594-604.
- Aslin, R. N. (2007). What's in a look? *Developmental Science*, 10, 48-53.
- Bailey, T., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E. and Rutter, M. (1995). Autism as a strongly genetic disorder: evidence from a British twin study. *Psychological Medicine*, 25, 63-77.
- Baillargeon, R. (1987). Object permanence in 3 ½- and 4 ½-month-old infants. *Developmental Psychology*, 23, 655-664.
- Baillargeon, R., Spelke, E. S. and Wasserman, S. (1985). Object permanence in five-month-old infants. *Cognition*, 20, 191-208.

- Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S. et al. (2000). A screening instrument for autism at 18 month of age: a six-year followup study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 694-702.
- Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., et al. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: The Special Needs and Autism Project (SNAP). *Lancet*, 368, 210–215.
- Baker, N. and Nelson, K. E. (1984). Recasting and related conversational techniques for triggering syntactic advances in young children. *First Language*, 5, 3-22.
- Baldwin, D.A. (1993). Early referential understanding: infants' ability to recognize referential acts for what they are. *Developmental Psychology*, 29, 832–843.
- Baldwin, D. A. (1995). *Understanding the link between joint attention and language*. In C. Moore and P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ: Erlbaum.
- Barnett, J. H., Croudace, T. J., Jaycock, S., Blackwell, C., Hynes, F., Sahakian, B. J., et al. (2007). Improvement and decline of cognitive function in schizophrenia over one year: a longitudinal investigation using latent growth modeling. *BioMed Central Psychiatry*, 7, 16.
- Baron-Cohen, S. (1989a). Joint-attention deficits in autism: Towards a cognitive analysis. *Development and Psychopathology*, 1, 185-189.
- Baron-Cohen, S. (1989b). Perceptual role taking and protodeclarative pointing in autism. *British Journal of Developmental Psychology*, 7, 113–127.

- Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Boston: MIT Press/Bradford Books.
- Baron-Cohen, S., Baldwin, D. A. and Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development*, 68, 48–57.
- Baron-Cohen, S., Jolliffe, T., Mortimore, C. and Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high-functioning adults with autism or Asperger Syndrome. *Journal of Child Psychology and Psychiatry*, 38, 813-822.
- Baron-Cohen, S., Leslie, A. M. and Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21, 37-46.
- Barresi, J. and Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, 19, 107-154.
- Bates, E. (1993). *Comprehension and production in early language development*. In E.S. Savage-Rumbaugh, J. Murphy, R.A. Sevcik, K.E. Brakke, S.L. Williams, and D.M. Rumbaugh (Eds.), *Monographs of the Society for Research in Child Development*, 58 (3–4, Serial No. 233), 222–242.
- Batki, A., Baron-Cohen, S., Wheelwright, S., Connellan, J. and Ahluwalia, J. (2000). Is there an innate gaze module? Evidence from human neonates. *Infant Behavior and Development*, 23, 223–229.
- Belmonte, M. K., Allen, G., Beckel-Mitchener, A., Boulanger, L. M., Carper, R. A. and Webb, S. J. (2004b). Autism and Abnormal Development of Brain Connectivity. *The Journal of Neuroscience*, 24, 9228-9231.

- Belmonte M. K., Cook Jr E. H., Anderson G. M., Rubenstein J. L. R., Greenough W. T., Beckel-Mitchener A., et al. (2004a). Autism as a disorder of neural information processing: directions for research and targets for therapy. *Molecular Psychiatry*, 9, 646-663.
- Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, 31, 419-456.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological Bulletin*, 112, 400-404.
- Bertone, A., Mottron, L., Jelenic, P. and Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain*, 128, 2430-2441.
- Berument, S. K., Rutter, M., Lord, C., Pickles, A. and Bailey, A. (1999). Autism screening questionnaire: diagnostic validity. *British Journal of Psychiatry*, 175, 444-451.
- Bijleveld, C. C. J. H. and van der Kamp, L. J.Th. (with Mooijaart, A., van der Kloot, W. A., van der Leeden, R., and van der Burg, E.). (1998). *Longitudinal data analysis: Designs, models and methods*. London: Sage.
- Blair, R. J. R. and Cipolotti, L. (2000). Impaired social response reversal: A case of acquired 'sociopathy'. *Brain*, 123, 1122-1141.
- Blalock, H. M. (1964). *Causal inferences in nonexperimental research*. Chapel Hill, NC: University of North Carolina.

- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., et al. (1994). A case-control family study of autism. *Journal of Child Psychology and Psychiatry*, 35, 877–900.
- Bomstein, M. H. and Sigman, M. D. (1986). Continuity in mental development from infancy. *Child Development*, 57, 251-274.
- British Psychological Society - Developmental Section, Conference Proceedings*. (2011) Newcastle, September.
- Brock, J., Brown, C. C., Boucher, J. and Rippon, G. (2002). The temporal binding deficit hypothesis of autism. *Developmental Psychopathology*, 14, 209-224.
- Brooks, R. and Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology*, 38, 958–966.
- Brooks, R. and Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science*, 8, 535–543.
- Browne, M. W. and Cudeck, R. (1993). *Alternative ways of assessing model fit*. In K. A. Bollen and J. S. Long (Eds.). *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Bryk, A. S. and Raudenbush, S. W. (1987). Applications of hierarchical linear models to assessing change. *Psychological Bulletin*, 101, 147-158.
- Bryson, S. E., Zwaigenbaum, L. McDermott, C., Rombough, V. and Brian, J. (2008). The Autism Observation Scale for Infants: Scale Development and Reliability Data. *Journal of Autism and Developmental Disorders*, 38, 731-738.

- Butterworth, G. and Jarrett, N. (1991). What minds have in common in space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9, 55–72.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus: Basic concepts, applications and programming*. New York: Routledge.
- Carey, S. and Bartlett, E. (1978). Acquiring a single new word. Paper presented at the Stanford Child Language Conference.
- Caron, A. J., Caron, R., Roberts, J. and Brooks, R. (1997). Infant sensitivity to deviations in dynamic facial-vocal displays: The role of eye regard. *Developmental Psychology*, 33, 802–813.
- Carpendale, J. I. M. and Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, 27, 79-151.
- Carpenter, M., Nagell, K. and Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, 63 (4, Serial No. 255).
- Carroll, J. M. and Snowling, M. J. (2004). Language and phonological skills in children at high risk of reading difficulties. *Journal of Child Psychology and Psychiatry*, 45, 631–640.
- Casey, B. J., Gordon, C. T., Mannheim, G. B. and Rumsey, J. M. (1993). Dysfunctional attention in autistic savants. *Journal of Clinical and Experimental Neuropsychology*, 15, 933–946.

- Cassel, T. D., Messinger, D. S., Ibanez, L. V., Haltigan, J. D., Acosta, S. I. and Buchman, A. C. (2007). Early social and emotional communication in the infant siblings of children with autism spectrum disorders: An examination of the broad phenotype. *Journal of Autism and Developmental Disorders*, 37, 122–132.
- Chang K. D., Steiner H. and Ketter T. A. (2000). Psychiatric phenomenology of child and adolescent bipolar offspring. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 453–460.
- Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society B*, 358, 315–324.
- Charman, T. and Baird, G. (2002). Practitioner review: Assessment and diagnosis of autism spectrum disorders in the pre-school years. *Journal of Child Psychology and Psychiatry*, 43, 289–305.
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Cox, A. and Drew, A. (2000). Testing joint attention, imitation, and play infancy precursors to language and theory of mind. *Cognitive Development*, 15, 481–498.
- Charman, T., Baron-Cohen, S., Swettenham, J., Baird, G., Drew, A. and Cox, A. (2003). Predicting language outcome in infants with autism and pervasive developmental disorder. *International Journal of Language and Communication Disorders*, 38, 265–285.
- Charman, T., Drew, A., Baird, C. and Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language*, 30, 213–236.

- Chawarska, K., Klin, A. and Volkmar, F. (2003). Automatic attention cueing through eye movement in 2-year-old children with autism. *Child Development*, 74, 1108–1122.
- Chawarska, K., Volkmar, F. and Klin, A. (2010). Limited attentional bias for faces in toddlers with autism spectrum disorders. *Archives of General Psychiatry*, 67, 178-185.
- Chevallier, C., Kohls, G., Troiani, V., Brodtkin, E. S. and Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, 16, 231-239.
- Christensen, L., Hutman, T., Rozga, A., Young, G. S., Ozonoff, S., Rogers, S. J., et al. (2010). Play and developmental outcomes in infant siblings of children with autism. *Journal of Autism and Developmental Disorders*, 40, 946-957.
- Clancy, B., Darlington, R. B. and Finlay, B. L. (2000). The course of human events: predicting the timing of primate neural development. *Developmental Science*, 3, 57–66.
- Colombo, J. (2002). Infant attention grows up: The emergence of a developmental cognitive neuroscience perspective. *Current Directions in Psychological Science*, 11, 196-200.
- Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N. and Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 45, 719–726.
- Constantino, J. N. and Todd, R. D. (2003). Autistic Traits in the General Population: A Twin Study. *Archives of General Psychiatry*, 60, 524-530.
- Constantino, J. N., Zhang, Y., Frazier, T., Abbacchi, A. M. and Law, P. (2010). Sibling recurrence and the genetic epidemiology of autism. *American Journal of Psychiatry*, 167, 1349–1356.

Courage, M. L., Howe, M. L. and Squires, S. E. (2004). Individual differences in 3.5-month olds' visual attention: What do they predict at 1 year? *Infant Behavior and Development*, 27, 19–30.

Courchesne, E. (1997). Brainstem, cerebellar and limbic neuroanatomical abnormalities in autism. *Current Opinion in Neurobiology*, 7, 269-278.

Courchesne, E. (2004). Brain development in autism: Early overgrowth followed by premature arrest of growth. *Mental Retardation and Developmental Disabilities Research Reviews. Special Issue: Epilepsy and the Development of Autism*, 10, 106-111.

Courchesne, E., Karns, C. M., Davis, H. R., Ziccardi, R., Carper, R. A., Tigue, Z. D., et al., (2001). Unusual brain growth patterns in early life in patients with autistic disorder. *Neurology*, 57, 245–254.

Csibra, G. and Gergely, G. (2006) *Social learning and social cognition: the case for pedagogy*. In Munakata, Y. and Johnson, M. H. (Eds), *Processes of Change in Brain and Cognitive Development. Attention and Performance XXI*. Oxford University Press, Oxford, pp. 249–274.

Csibra, G., Tucker, L. A. and Johnson, M. H. (1998). Neural correlates of saccade planning in infants: A high-density ERP study. *International Journal of Psychophysiology*, 29, 201-125.

Csibra, G., Tucker, L. A. and Johnson, M. H. (2001). Differential frontal cortex activation before anticipatory and reactive saccades in infants. *Infancy*, 2, 159-174.

Curran, P. J. and Hussong, A. M. (2003). The use of latent trajectory models in psychopathology research. *Journal of Abnormal Psychology*, 112, 526-544.

- Dale, P. S. and Fenson, L. (1996). Lexical development norms for young children. *Behavioral Research Methods, Instruments, and Computers*, 28, 125-127.
- Davidson, D., Jergovic, D., Imami, Z. and Theodos, V. (1997). Monolingual and bilingual children's use of the mutual exclusivity constraint. *Journal of Child Language*, 24, 3-24.
- Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J. and Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, 28, 479-485.
- Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., et al. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental Psychology*, 40, 271-283.
- de Haan, M., Pascalis, O. and Johnson, M. H. (2002). Specialization of neural mechanisms underlying face recognition in human infants. *Journal of Cognitive Neuroscience*, 14, 199-209.
- de Marchena, A., Eigsti, I. M., Worek, A., Ono, K. E. and Snedeker, J. (2011). Mutual exclusivity in autism spectrum disorders: testing the pragmatic hypothesis. *Cognition*, 119, 96-113.
- Deaner, R. O. and Platt, M. L. (2003). Reflexive social attention in monkeys and humans. *Current Biology*, 13, 1609-1613.
- Dichter, G. S., Radonovich, K. J., Turner-Brown, L. M., Lam, K. S., Holtzclaw, T. N. and Bodfish, J. W. (2009). Performance of children with autism spectrum disorders on the dimension-change card sort task. *Journal of Autism and Developmental Disorders*, 40, 448-456.

- Dichter, G. S., Richey, J. A., Rittenberg, A. M., Sabatino, A. and Bodfish, J. W. (2012). Reward circuitry function in autism during face anticipation and outcomes. *Journal of Autism and Developmental Disorders*, 42, 147–160.
- DiLalla, D. L. and Rogers, S. J. (1994). Domains of the Childhood Autism Rating Scale: relevance for diagnosis and treatment. *Journal of Autism and Developmental Disorders*, 24, 115–128.
- Dimigen, O., Sommer, W., Hohlfeld, A., Jacobs, A. and Kliegl, R. (2011). Coregistration of eye movements and EEG in natural reading: Analyses and review. *Journal of Experimental Psychology: General*, 140, 552-572.
- Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E. and Baron- Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. *Visual Cognition*, 6, 509–540.
- Dunn, G., Everitt, B. and Pickles, A. (1993). *Modelling covariances and latent variables using EQS*. Chapman & Hall, London.
- Eigsti, I-M, de Marchena, A. B., Schuh, J. M. and Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, 5, 681-691.
- Elsabbagh, M. and Johnson, M. (2007). Infancy and autism: progress, prospects, and challenges. *Progress in Brain Research*, 164, 355-383.
- Elsabbagh, M. and Johnson, M. H. (2010). Getting answers from babies about autism. *Trends in Cognitive Sciences*, 14, 81-87.

Elsabbagh, M., Bedford, R., Senju, A., Charman, T., Pickles, A., Johnson, M. H. et al. (under review). What you see is what you get: Endogenous and exogenous influences on face scanning in typical and atypical development.

Elsabbagh, M., Fernandes, J., Webb, S. J., Dawson, G., Charman, T. and Johnson, M. H. (under review). Disengagement of visual attention at 14-months of age is associated with emerging autism in toddlerhood.

Elsabbagh, M., Mercure, E., Hudry, K., Chandler, S., Pasco, G., Charman, T., et al. (2012). Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology*, 22, 1-5.

Elsabbagh, M., Volein, Á., Csibra, G., Holmboe, K., Garwood, H., Tucker, L., et al. (2009b). Neural correlates of eye gaze processing in the infant broader autism phenotype. *Biological Psychiatry*, 65, 31–38.

Elsabbagh, M., Volein, Á., Holmboe, K., Tucker, L., Csibra, G., Baron-Cohen, et al. (2009a). Visual orienting in the early broader autism phenotype: disengagement and facilitation. *Journal of Child Psychology and Psychiatry*, 50, 637–642.

Erlenmeyer-Kimling, L. (2000). Neurobehavioral deficits in offspring of schizophrenic parents: Liability indicators and predictors of illness. *American Journal of Medical Genetics*, 97, 65–71.

Fagan, J. F. (1984). The Intelligent Infant: Theoretical Implications. *Intelligence*, 8, 1-9.

Fagan, J. F. and Montie, J. E. (1988). Racial differences in IQ: Item analysis of the Stanford-Binet at 3 years. *Intelligence*, 12, 315-332.

- Faraone, S. V., Biedernam, J., Mennin, D., Gershon, J., Tsuang, M. T. (1996). A prospective four-year follow-up study of children at risk for ADHD: psychiatric, neuropsychological, and psychosocial outcome. *Journal of the American Academy of Child and Adolescent Psychiatry*, 33, 1449-1459.
- Farroni, T., Csibra, G., Simion, F. and Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 9602–9605.
- Farroni, T., Mansfield, E. M., Lai, C. and Johnson, M. H. (2003). Infants perceiving and acting on the eyes: Tests of an evolutionary hypothesis. *Journal of Experimental Child Psychology*, 85, 199–212.
- Farroni, T., Massaccesi, S., Pividori, D. and Johnson, M. H. (2004). Gaze following in Newborns. *Infancy*, 5, 39–60.
- Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P. et al. (1993). *MacArthur Communicative Development Inventories*. Baltimore: Paul H. Brookes Publishing Co.
- Fine, C., Lumsden, J. and Blair, R. J. R. (2001). Dissociation between ‘theory of mind’ and executive functions in a patient with early left amygdala damage. *Brain*, 124, 287-298.
- Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders*, 33, 365-382.
- Fombonne, E., Simmons, H., Ford, T., Meltzer, H. and Goodman, R. (2003). Prevalence of pervasive developmental disorders in the British nationwide survey of child mental health. *International Review of Psychiatry*, 15, 158–165.

- Fransson, P., Åden, U., Blennow, M. and Lagercrantz, H. (2011). The functional architecture of the infant brain as revealed by resting-state fMRI. *Cerebral Cortex*, 21, 145-154.
- Frischen, A., Bayliss, A. P. and Tipper, S. P. (2007). Gaze cueing of attention. Visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133, 694–724.
- Frith, U. (1989). *Autism: explaining the enigma*. Oxford: Blackwell.
- Frith, U. (2003). *Autism: Explaining the enigma* (2nd ed.). Oxford: Blackwell.
- Frith, U. and Happé, F. (1994). Autism: Beyond “theory of mind”. *Cognition*, 50, 115-132.
- Gao, W. and Lin, W. (2012). Frontal parietal control network regulates the anti-correlated default and dorsal attention networks. *Human Brain Mapping*, 33, 192-202.
- Garon, N., Bryson, S. E., Zwaigenbaum, L., Smith, I. M., Brian, J., Roberts, W., et al. (2009). Temperament and its relationship to autistic symptoms in a high-risk infant sib cohort. *Journal of Abnormal Child Psychology*, 37, 59-78.
- Gliga, T., Elsabbagh, M., Hudry, K., Charman, T., Johnson, M. H., and the BASIS team. (2012). Gaze following, gaze reading and word learning in children at-risk for autism. *Child Development*, 83, 926-938
- Goldberg, W. A., Jarvis, K. L., Osann, K., Laulhere, T. M., Straub, C., Thomas, E., et al. (2005). Brief report: Early social communication behaviors in the younger siblings of children with autism. *Journal of Autism and Developmental Disorders*, 35, 657–664.
- Goldstein, G., Johnson, C. R. and Minshew, N. J. (2001). Attentional processes in autism. *Journal of Autism and Developmental Disorders*, 31, 433-440.

Goodale, M. A. and Milner, A.D. (1992). Separate visual pathways for perception and action.

Trends in Cognitive Sciences, 15, 20–25.

Goodman, R., Ford, T., Richards, H., Gatward, R. and Meltzer, H. (2000). The development and well-being assessment: Description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, 41, 645–655.

Green, J., Charman, T., McConachie, H., Aldred, C., Slonims, V., Howlin, P. et al. (2010). Parent-mediated communication-focused treatment in children with autism (PACT): A randomised controlled trial. *Lancet*, 375, 2152–2160.

Grice, S. J., Halit, H., Farroni, T., Baron–Cohen, S., Bolton, P. and Johnson, M. H. (2005). Neural correlates of eye-gaze detection in young children with autism. *Cortex*, 41, 342–353.

Gross, H. E., Shaw, D. S., Burwell, R. A. and Nagin, D. S. (2009). Transactional processes in child disruptive behavior and maternal depression: A longitudinal study from early childhood to adolescence. *Development and Psychopathology*, 21, 139–156.

Hadjikhani, N., Joseph, R. M., Snyder, J. and Tager-Flusberg, H. (2007). Abnormal activation of the social brain during face perception in autism. *Human Brain Mapping*, 28, 441–449.

Haith, M. M. (1998). Who put the cog in infant cognition? Is rich interpretation too costly? *Infant Behavior and Development*, 21, 167–179.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, 87, B23–34.

- Halit, H., Csibra, G., Volein, Á. and Johnson, M. H. (2004). Face-sensitive cortical processing in early infancy. *Journal of Child Psychology and Psychiatry*, 45, 1228–1234.
- Happé, F. and Frith, U. (2006). The weak coherence account: detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 36, 5–25.
- Happé, F. and Roland, A. (2008). ‘Fractionable Autism Triad’: A Review of Evidence from Behavioural, Genetic, Cognitive and Neural Research. *Neuropsychology Review*, 18, 287-304.
- Hare, B., Call, J. and Tomasello, M. (1998). Communication of food location between human and dog (*Canis Familiaris*). *Evolution of Communication*, 2, 137–159.
- Hill, E. L. (2004). Executive dysfunction in autism. *Trends in Cognitive Sciences*, 8, 26–32.
- Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., et al. (2000). Breaking the language barrier: an emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development*, 65, i-vi, 1-123.
- Holmboe, K., Elsabbagh, M., Volein, Á., and Tucker, L., Baron-Cohen, S., Bolton, et al. (2010). Frontal cortex functioning in the infant broader autism phenotype. *Infant Behaviour and Development*, 33, 482-491.
- Hood, B. M., and Atkinson, J. (1993). Disengaging visual-attention in the infant and adult. *Infant Behavior and Development*, 16, 423–439.

- Hood, B. M., Willen, J. D. and Driver, J. (1998). Gaze perception triggers corresponding shifts of visual attention in young infants. *Psychological Science*, 9, 131–134.
- Horst, J. S. and Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, 13, 128-157.
- Horst, J. S., Scott, E. J. and Pollard, J. A. (2010). The role of competition in word learning via referent selection. *Developmental Science*, 13, 706-713.
- Houston-Price, C., Plunkett, K. and Duffy, H. (2006). The use of social and salience cues in early word learning. *Journal of Experimental Child Psychology*, 95, 27-55.
- Hu, L-T. and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H. Aldred, C., et al. (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International Journal of Language and Communication Disorders*, 45, 681-690.
- Hunt, J. M. (1961). *Intelligence and experience*. Oxford, England: Ronald
- Hutman, T., Rozga, A., DeLaurentis, A. D., Barnwell, J. M., Sugar, C. A. and Sigman, M. (2010). Response to distress in infants at risk for autism: A prospective longitudinal study. *Journal of Child Psychology and Psychiatry*, 51, 1010-1020.
- Johnson, M. H. (1990). Cortical Maturation and the Development of Visual Attention in Early Infancy. *Journal of Cognitive Neuroscience*, 2, 81-95

- Johnson, M.H., 2001. Functional brain development in humans. *Nature Reviews Neuroscience*, 2, 475–483.
- Johnson, M. H. (2011). Interactive specialization: a domain-general framework for human functional brain development? *Developmental Cognitive Neuroscience*, 1, 7-21
- Johnson, M. H., Griffin, R., Csibra, G., Halit, H., Farroni, T., de Haan, M. et al. (2005). The emergence of the social brain network: Evidence from typical and atypical development. *Developmental Psychopathology*, 17, 599-619.
- Johnson, M. H., Posner, M. I. and Rothbart, M. K. (1991). Components of visual orienting in early infancy: Contingency learning, anticipatory looking, and disengaging. *Journal of Cognitive Neuroscience*, 3, 335-344.
- Just, M. A., Cherkassky, V. L., Keller, T. A. and Minshew, N. J. Cortical activation and synchronization during sentence comprehension in high-functioning autism: Evidence of underconnectivity. *Brain*, 2004, 127, 1811–21.
- Jöreskog, K. G. (1993). *Testing structural equation models*. In K. A. Bollen and J. S. Long (Eds.). *Testing structural equation models* (pp. 294-316). Newbury Park, CA: Sage.
- Jöreskog, K. G. and Sörbom, D. (1993). *LISREL 8: Structural equation modelling with the SIMPLIS command language*. Chicago: Scientific Software International.
- Kang, S. H., Pashler, H., Cepeda, N. J., Rohrer, D., Carpenter, S. K. and Mozer, M. C. (2011). Does incorrect guessing impair fact learning? *Journal of Educational Psychology*, 103, 48-59.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217-250.

- Karmiloff-Smith, A. (2006). The tortuous route from genes to behavior: A neuroconstructivist approach. *Cognitive, Affective and Behavioral Neuroscience*, 6, 9-17.
- Kasari, C., Paparella, T., Freeman, S. and Jahromi, L. B. (2008). Language outcome in autism: randomized comparison of joint attention and play interventions. *Journal of Consulting and Clinical Psychology*, 76, 125-137.
- Kerner, B. and Muthén, B. (2009). Growth mixture modelling in families of the Framingham Heart Study. *BioMed Central Proceedings*, 3, 1-5.
- Kleinmans, N., Akshoomoff, N. and Delis, D. C. (2005). Executive functions in autism and Asperger's disorder: flexibility, fluency, and inhibition. *Developmental Neuropsychology*, 27, 379-401.
- Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G. and Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, 459, 257-261.
- Kramer, J. H. and Quitania, L. (2007). *Bedside frontal lobe testing*. In B. L. Miller and J. L. Cummings (Eds.), *The human frontal lobes: Functions and disorders* (pp. 279-291). (2nd ed.). New York: Guilford.
- Landa, R. and Garrett-Mayer, E. (2006). Development in infants with autism spectrum disorders: a prospective study. *Journal of Child Psychology and Psychiatry*, 47, 629-638.
- Landa, R. J., Holman, K. C. and Garrett-Mayer, E. (2007). Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Archives of General Psychiatry*, 64, 853-864.

- Landry, R. and Bryson, S. E. (2004). Impaired disengagement of attention in young children with autism. *Journal of Child Psychology and Psychiatry*, 45, 1115-1122.
- Lecavalier, L. (2005). An evaluation of the Gilliam Autism Rating Scale. *Journal of Autism and Developmental Disorders*, 35, 795–805.
- Leekam, S., López, B. and Moore, C. (2000). Attention and joint attention in preschool children with autism. *Developmental Psychology*, 36, 261-273.
- Little, R. J. A. and Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Little, R. J. A. and Rubin, D. B. (1989). The analysis of social science data with missing values. *Sociological Methods and Research*, 18, 292-326.
- Loh, A., Soman, T., Brian, J., Bryson, S. E., Roberts, W., Szatmari, P., et al. (2007). Stereotyped motor behaviors associated with autism in high-risk infants: A pilot videotape analysis of a sibling sample. *Journal of Autism and Developmental Disorders*, 37, 25-36.
- Lopez, B. R., Lincoln, A. J., Ozonoff, S. and Lai, Z. (2005). Examining the relationship between executive functions and restricted, repetitive symptoms of autistic disorder. *Journal of Autism and Developmental Disorders*, 35, 445–460.
- Lord, C., Risi, S. and Pickles, A. (2004). *Trajectory of language development in autism spectrum disorders*. In M. Rice and S. Warren (Eds.), *Developmental language disorders: From phenotypes to etiologies* (pp. 7–29). Mahwah, NJ: Lawrence Erlbaum.

- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communicative deficits associated with the spectrum of autism. *Journal of Autism Developmental Disorders*, 30, 205–223.
- Lord, C., Rutter, M. L. and LeCouteur, A. (1994). The autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659–685.
- Losh, M. and Piven, J. (2007). Social-cognition and the broad autism phenotype: identifying genetically meaningful phenotypes. *Journal of Child Psychology and Psychiatry*, 48, 105–112.
- Luna, B., Thulborn, K.R., Munoz, P.D., Merriam, E.P., Garver, K.E., Minshew, N.J., Keshavan, M.S., Genovese, C.R., Eddy, W.F., Sweeney, J.A., 2001. Maturation of widely distributed brain function subserves cognitive development. *NeuroImage*, 13, 786–793.
- MacCallum, R. C., Browne, M. W. and Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, 1, 130–149.
- MacWhinney, B. (1989). *Competition and lexical categorization*. In R. Corrigan, F. Eckman and M. Noonan (Eds.). *Linguistic categorization*, (pp. 195–241). Amsterdam, the Netherlands: John Benjamin.

- Markman, E. M. (1991). *The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings*. In S. A. Gelman and P. Byrnes (Eds.), *Perspectives on Language and Thought* (pp. 80).
- Markman, E. M. and Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20, 121-157.
- Markman, E. M. Wasow, J. L. and Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, 47, 241-275.
- Mather, E. and Plunkett, K. (2010). Novel labels support 10-month-olds' attention to novel objects. *Journal of Experimental Child Psychology*, 105, 232-242.
- Mather, E. and Plunkett, K. (2011). Mutual exclusivity and phonological novelty constrain word learning at 16 months. *Journal of Child Language*, 38, 933-950.
- McCall, R. B. (1994). What process mediates predictions of childhood IQ from infant habituation and recognition memory? Speculations on the roles of inhibition and rate of information processing. *Intelligence*, 18, 107-125.
- McCall, R. B. and Carriger, M. (1993). A meta-analysis of infant habituation and recognition memory performance as predictors of later IQ. *Child Development*, 64, 57-79.
- McConnell, B.A. and Bryson, S.E. (2005). Visual attention and temperament: Developmental data from the first 6 months of life. *Infant Behavior and Development*, 28, 537-544.
- McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, 317, 631.
- McNelis N. L. and Boatright-Horowitz S. L. (1998). Social monitoring in a primate group: the relationship between visual attention and hierarchical ranks. *Animal Cognition*, 1, 65-69.

- Merriman, W. E. and Schuster, J. M. (1991). Young children's disambiguation of object name reference. *Child Development*, 62, 1288-1301.
- Miranda-Linne, F. M. and Melin, L. (2002). A factor analytic study of the Autism Behavior Checklist. *Journal of Autism and Developmental Disorders*, 32, 181–188.
- Moll, H. and Tomasello, M. (2010). Infant cognition. *Current Biology*, 20, R872-R875.
- Moore, C. and Corkum, V. L. (1994). Social understanding at the end of the first year of life. *Developmental Review*, 14, 349–372.
- Moore, C. and Corkum, V. (1998). Infant gaze following based on eye direction. *British Journal of Developmental Psychology*, 16, 495–503.
- Morales, M., Mundy, P., Delgado, C. E. F., Yale, M., Messinger, D., Neal, R. et al. (2000). Responding to joint attention across the 6- through 24-month age period and early language acquisition. *Journal of Applied Developmental Psychology*, 21, 283-298.
- Morales, M., Mundy, P. and Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development*, 21, 373–377.
- Morton, J., and Johnson, M. H. (1991). CONSPEC and CONLEARN: A two-process theory of infant face recognition. *Psychological Review*, 98, 164-181.
- Mottron, L., and Burack, J.A. (2001). *Enhanced perceptual functioning in the development of autism*. In J.A. Burack, T. Charman, N. Yirmiya, and P.R. Zelazo (Eds.), *The development of autism: Perspectives from theory and research* (pp. 131–148). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mottron, L., Burack, J. A., Iarocci, G., Belleville, S. and Enns, J. T. (2003). Locally oriented perception with intact global processing among adolescents with high-functioning

- autism: evidence from multiple paradigms. *Journal of Child Psychology and Psychiatry*, 44, 904–13.
- Mottron, L., Burack, J., Stauder, J. and Robaey, P. (1999). Perceptual processing among high-functioning persons with autism. *Journal of Child Psychology and Psychiatry*, 40, 203–211.
- Mullen, E. M. (1995). *Mullen Scales of Early Learning* (AGS ed.). Circle Pines, MN: American Guidance Service Inc.
- Mundy, P., Card, J., Fox, N. (2000). Fourteen-month cortical activity and different infant joint attention skills. *Developmental Psychobiology*, 36, 325–338.
- Mundy, P., Hogan, A. and Doehring, P. (1996). *A preliminary manual for the abridged early social-communication scales*. Coral Gables, FL: University of Miami.
www.psy.miami.edu/faculty/pmundy.
- Mundy, P. and Neal, R. (2001). *Neural plasticity, joint attention and a transactional social-orienting model of autism*. In L. Glidden (Ed.), *International review of research in mental retardation*: Vol. 23. Autism (pp. 139–168). New York: Academic Press.
- Mundy, P. and Newell, L. (2007). Attention, Joint Attention, and Social Cognition. *Current Directions in Psychological Science*, 16, 269-274.
- Mundy, P., Sigman, M., Ungerer, J. and Sherman, T. (1986). Defining the social deficits of autism: the contribution of non-verbal communication measures. *Journal of Child Psychology and Psychiatry*, 27, 657-669.

- Mundy, P., Sullivan, L. and Mastergeorge, A. M. (2009). A parallel and distributed-processing model of joint attention, social cognition and autism. *Autism Research*, 2, 2–21.
- Muthén, B. and Khoo, S. T. (1998). Longitudinal studies of achievement growth using latent variable modeling. *Learning and Individual Differences, Special issue: latent growth curve analysis*, 10, 73-101.
- Nadig, A. S., Ozonoff, S., Young, G. S., Rozga, A., Sigman, M. and Rogers, S. J. (2007). A prospective study of response to name in infants at risk for autism. *Archives of Pediatric and Adolescent Medicine*, 161, 378–83.
- Nation, K. and Penny, S. (2008). Sensitivity to eye gaze in autism: is it normal? Is it automatic? Is it social? *Development and Psychopathology*, 20, 79-97.
- Osterling, J., and Dawson, G. (1994). Early, recognition of children with autism: A study of first birthday home video tapes. *Journal of Autism and Developmental Disorders*, 24, 247—257.
- Ozonoff, S., Iosif, A-M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., et al. (2010) A Prospective Study of the Emergence of Early Behavioral Signs of Autism. *Journal of the American Academy of Child and Adolescent Psychiatry*, 49, 256-266.
- Ozonoff, S., Macari, S., Young, G. S., Goldring, S., Thompson, M. and Rogers, S. J. (2008). Atypical object exploration at 12 months of age is associated with autism in a prospective sample. *Autism*, 12, 457-472.
- Ozonoff, S., Young, G. S., Carter, A., Messinger, D., Yirmiya, N., Zwaigenbaum, L., et al. (2011). Recurrence risk for autism spectrum disorders: A baby siblings research consortium study. *Pediatrics*, 128, e1–e8.

- Parish-Morris, J., Hennon, E. A., Hirsh-Pasek, K., Golinkoff, R. M. and Tager-Flusberg, H. (2007). Children with autism illuminate the role of social intention in word learning. *Child Development*, 78, 1265-1287.
- Pellicano, E. (2007). Links Between Theory of Mind and Executive Function in Young Children With Autism: Clues to Developmental Primacy. *Developmental Psychology*, 43, 974-990.
- Pelphrey, K. A. and Perlman, S. B. (2009). *Charting brain mechanisms for the development of social cognition*. In J. M. Rumsey and M. Ernst (Eds.), *Neuroimaging in Developmental Clinical Neuroscience*. Cambridge University Press.
- Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.
- Piaget, J. (1954). *The construction of reality in the child*. New York: Basic Books.
- Pickles, A. (1998). *Generalized estimating equations*. In P. Armitage, T. Colton (Eds.), *The encyclopedia of biostatistics Vol 2* (pp. 1626–1637). New York: Wiley.
- Plaisted, K., Swettenham, J. and Rees, L. (1999). Children with autism show local precedence in a divided attention task and global precedence in a selective attention task. *Journal of Child Psychology and Psychiatry*, 40, 733–742.
- Poljac, E., Simon, S., Ringlever, L., Kalcik, D., Groen, W. B., Buitelaar, J. K., et al. (2010). Impaired task switching performance in children with dyslexia but not in children with autism. *Quarterly Journal of Experimental Psychology (Colchester)*, 63, 401-416.

- Povinelli, D. J., Bierschwale, D. T. and Cech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. *British Journal of Developmental Psychology*, 17, 37–60.
- Preissler, M. A. and Carey, S. (2005). The role of inferences about referential intent in word learning: evidence from autism. *Cognition*, 97, B13-23.
- Presmanes, A. G., Walden, T. A., Stone, W. L. and Yoder, P. J. (2007). Effects of different attentional cues on responding to joint attention in younger siblings of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 37, 133-144.
- Quine, W. V. O. (1960). *Word and object*. Cambridge MA: MIT Press.
- Raftery, A. E. (1993). *Bayesian model selection in structural equation models*. In K. A. Bollen and J. S. Long (Eds.). *Testing structural equation models* (pp. 163-180). Newbury Park, CA: Sage.
- Rindskopf, D. (1984). Structural equation models: Empirical identification, Heywood cases and related problems. *Sociological Methods and Research*, 13, 109-119.
- Risi, S., Lord, C., Gotham, K., Corsello, C., Chrysler, C., Szatmari, P., et al. (2006). Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45, 1094–1103.
- Rogers, S. J. (2009). What are infant siblings teaching us about autism in infancy? *Autism Research*, 2, 125–137.

Ronald, A., Happé, F., Bolton, P., Butcher, L. M., Price, T. S., Wheelwright, S., et al. (2006).

Genetic heterogeneity between the three components of the autism spectrum: a twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 45, 691–699.

Ronald, A., Happé, F. and Plomin, R. (2005). The genetic relationship between individual differences in social and non-social behaviours characteristic of autism.

Developmental Science, 8, 444–458.

Rothbart, M. K., Posner, M. I. and Rosicky, J. (1994). Orienting in normal and pathological development. *Development and Psychopathology*, 6, 635–652.

Rubenstein, J. L. and Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes Brain and Behavior*, 2, 255–267.

Russell, J. and Hill, E. L. (2001). Action-monitoring and intention reporting in children with autism. *Journal of Child Psychology and Psychiatry*, 42, 317–328.

Russell, J., Mauthner, N., Sharpe, S. and Tidswell, T. (1991). The “windows task” as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology*, 9, 331–349.

Russell, J., Saltmarsh, R. and Hill, E. (1999). What do executive factors contribute to the failure on false belief tasks by children with autism? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 40, 859–868.

Rutter, M., Bailey, A. and Lord, C. (2003) *SCQ: Social communication questionnaire*. Manual. Los Angeles, CA: Western Psychological Services.

- Saffran, J. R., Aslin, R. N. and Newport, E. L. (2008). Statistical learning by 8-month-old infants. *Science*, 274, 1926-1928.
- Samuelson, L. and Smith, L. (1998). Memory and attention make smart word learning: an alternative account of Akhtar, Carpenter, and Tomasello. *Child Development*, 69, 94-104.
- Satorra, A. and Saris, W. E. (1985). Power of the likelihood ratio test. *Psychometrika*, 50, 83-90.
- Scaife, M. and Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, 253, 265-266.
- Schubert, E. W. and McNeil, T. F. (2004). Prospective study of neurological abnormalities in offspring of women with psychosis: Birth to adulthood. *American Journal of Psychiatry*, 161, 1030-1037.
- Schwartz, G. (1978). Estimating the dimension of a model. *Annals of statistics*, 6, 461-464.
- Scott-Van Zeeland, A. A. (2010). Reward processing in autism. *Autism Research*, 3, 53-67
- Senju, A. and Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18, 668-671.
- Senju, A., Tojo, Y., Dairoku, H. and Hasegawa, T. (2004). Reflexive orienting in response to eye gaze and an arrow in children with and without autism. *Journal of Child Psychology and Psychiatry*, 45, 445-458.
- Shah, A. and Frith, U. J. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry*, 34, 1353-1364.

- Sigman, M. and Ruskin, E. (1999). *Social competence in children with autism, Down syndrome and developmental delay: A longitudinal study*. Monograph of the Society for Research in Child Development. Serial No. 256, 64(1), Chicago, IL: University of Chicago Press.
- Siller, M. and Sigman, M. (2008). Modeling longitudinal change in the language abilities of children with autism: parent behaviors and child characteristics as predictors of change. *Developmental Psychology*, 44, 1691-1704.
- Simon, H. A. (1954). Spurious correlation: A causal interpretation. *Journal of the American Statistical Association*, 49, 922-932.
- Southgate, V., Senju, A. and Csibra, G. (2007). Action anticipation through attribution of false belief in two-year-olds. *Psychological Science*, 18, 587- 592.
- Spencer, J., O'Brien, J., Riggs, K., Braddick, O., Atkinson, J. and Wattam-Bell, J. (2000). Motion processing in autism: evidence for a dorsal stream deficiency. *Neuroreport*, 11, 2765-2767.
- Stella, J., Mundy, P. and Tuchman, R. (1999). Social and non-social factors in the Childhood Autism Rating Scale. *Journal of Autism and Developmental Disorders*, 29, 307-317.
- Striano, T., Reid, V. M. and Hoehl, S. (2006). Neural mechanisms of joint attention in infancy. *European Journal of Neuroscience*, 23, 2819-2823.
- Stuss, D. T. (2007). *New approaches to prefrontal lobe testing*. In B. L. Miller and J. L. Cummings (Eds.), *The human frontal lobes: Functions and disorders* (pp. 292-305). (2nd ed.). New York: Guilford.

- Sullivan, M., Finelli, J., Marvin, A., Garrett-Mayer, E., Bauman, M. and Landa, R. (2007). Response to joint attention in toddlers at risk for autism spectrum disorder: A prospective study. *Journal for Autism and Developmental Disorders*, 37, 37–48.
- Swettenham, J., Condie, S., Campbell, R., Milne, E. and Coleman, M. (2003). Does the perception of moving eyes trigger reflexive visual orienting in autism? *Philosophical Transactions of the Royal Society B*, 358, 325–334.
- Szatmari, P., Merette, C., Bryson, S. E., Thivierge, J., Roy, M. A., Cayer, M., et al., (2002). Quantifying dimensions in autism: a factor-analytic study. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 467–474.
- Tabachnick, B. G. and Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson.
- Tadevosyan-Leyfer, O., Dowd, M., Mankoski, R., Winklosky, B., Putnam, S., McGrath, L., et al. (2003). A principal components analysis of the autism diagnostic interview—revised. *Journal of the American Academy of Child and Adolescent Psychiatry*, 42, 864–872.
- Tager-Flusberg, H., Paul, R. and Lord, C. E. (2005). *Language and communication in autism*. In: F. Volkmar, R. Paul, A. Klin, D. J. Cohen (Eds). *Handbook of autism and pervasive developmental disorder*. 3rd ed. Vol. 1. New York: Wiley; pp. 335–364.
- Tek, S., Jaffery, G., Fein, D. and Naigles, L. R. (2008). Do children with autism spectrum disorders show a shape bias in word learning? *Autism Research*, 1, 208–222.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

- Tomasello, M., Call, J. and Hare, B. (1998). Five primate species follow the visual gaze of conspecifics. *Animal Behavior*, 55, 1063–1069.
- Tomasello, M., Carpenter, M., Call, J., Behne, T. and Moll, H. (2005). Understanding sharing intentions: The origins of cultural cognition. *Brain and Behavior Sciences*, 28, 675–690.
- Tomasello, M. and Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 57, 1454–1463.
- Toth, K., Dawson, G., Meltzoff, A. N., Greenson, J. and Fein, D. (2007). Early social, imitation, play, and language abilities of young non-autistic siblings of children with autism. *Journal of Autism and Developmental Disorders*, 37, 145-157.
- Toth, K., Munson, J., Meltzoff, A. N. and Dawson, G. (2006). Early predictors of communication development in young children with autism spectrum disorder: Joint attention, imitation, and toy play. *Journal of Autism and Developmental Disorders*, 36, 993–1005.
- Tsotsos, J. K., Culhane, S. M., Yan Kei Wai, W., Lai, Y., Davis, N. and Nuflo, F. (1995). Modeling visual attention via selective tuning. *Artificial Intelligence*, 78, 507-545.
- Van Der Geest, J. N., Kemner, C., Camfferman, G., Verbaten, M. N. and Van Engeland, H. (2001). Eye movements, visual attention, and autism: A saccadic reaction time study using the gap and overlap paradigm. *Biological Psychiatry*, 50, 614–619.
- van Lang, N. D., Boomsma, A., Sytema, S., de Bildt, A. A., Kraijer, D. W., Ketelaars, C., et al. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *Journal of Child Psychology and Psychiatry*, 47, 37–44.

Vivanti, G., McCormick, C., Young, G. S., Abucayan, F., Hatt, N., Nadig, A., et al. (2011).

Intact and impaired mechanisms of action understanding in autism. *Developmental Psychology*, 47, 841–856.

Volkmar, F. R., Cicchetti, D. V., Dykens, E., Sparrow, S. S., Leckman, J. F. and Cohen, D. J.

(1988). An evaluation of the autism behavior checklist. *Journal of Autism and Developmental Disorders*, 18, 81–97.

Wadden, N. P. K., Bryson, S. E. and Ridger, R. S. (1991). A closer look at the autism

behavior checklist: discriminant validity and factor structure. *Journal of Autism and Developmental Disorders*, 21, 529–541.

Wass, S. V., Porayska-Pomsta, K. and Johnson, M. H. (2011). Training attentional control in

infancy. *Current Biology*, 21, 1543-1547.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio,

TX:Harcourt Assessment.

Williams, J. H., Waiter, G. D., Perra, O., Perrett, D. I. and Whiten, A. (2005). An fMRI study

of joint attention experience. *NeuroImage*, 25, 133–140.

Wing, L. and Wing, J. K. (1971). Multiple impairments in early childhood autism. *Journal of*

Autism and Developmental Disorders, 1, 256-266.

Witkin, H., Oltman, P., Raskin, E. and Karp, S. (1971). *A manual for the Embedded Figures*

Test. Palo Alto, CA: Consulting Psychologists Press.

Wolff, J. J., Gu, H., Gerig, G., Elison, J. T., Styner, M., Gouttard, S., Botteron, K. N., et al.

(2012). Differences in white matter fiber tract development present from 6 to 24 months in infants with autism. *American Journal of Psychiatry*, 169, 589-600.

- Woodward, A. L. and Hoyne, K. L. (1999). Infants' learning about words and sounds in relation to objects. *Child Development*, 70, 65-77.
- World Health Organisation (1993). *Mental Disorders: A glossary and guide to their classification in accordance with the 10th revision of the international classification of diseases: Research diagnostic criteria (ICD-10)*. Geneva: WHO.
- Yirmiya, N., and Charman, T. (2010). The prodrome of autism: Early behavioral and biological signs, regression, peri- and post-natal development and genetics. *Journal of Child Psychology and Psychiatry*, 51, 432–458.
- Yirmiya, N., Gamliel, I., Pilowsky, T., Feldman, R., Baron-Cohen, S. and Sigman, M. (2006). The development of siblings of children with autism at 4 and 14 months: Social engagement, communication, and cognition. *Journal of Child Psychology and Psychiatry*, 47, 511–523.
- Yirmiya, N., Gamliel, I., Shaked, M. and Sigman, M. (2007). Cognitive and verbal abilities of 24- to 36-month-old siblings of children with autism. *Journal of Autism and Developmental Disorders*, 37, 218-229.
- Yoder, P., Stone, W. L., Walden, T. and Malesa, E. (2009). Predicting social impairment and ASD diagnosis in younger siblings of children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 39, 1381-1391.
- Young, G. S., Merin, N., Rogers, S. J. and Ozonoff, S. (2009). Gaze Behavior and Affect at 6-Months: Predicting Clinical Outcomes and Language Development in Typically Developing Infants and Infants At-Risk for Autism. *Developmental science*, 12, 798–814.

Young, G. S., Rogers, S. J., Hutman, T., Rozga, A., Sigman, M. and Ozonoff, S. (2011).

Imitation from 12 to 24 months in autism and typical development: A longitudinal Rasch analysis. *Developmental Psychology*, 47, 1565-1578.

Zwaigenbaum, L., Bryson, S., Rogers, T., Roberts, W., Brian, J. and Szatmari, P. (2005).

Behavioral manifestations of autism in the first year of life. *International Journal of Developmental Neuroscience*, 23, 143-152.